

Prompted vs. Organic Customer Insight: Comparing the Value of Focus Groups and Online Reviews in Product and Service Innovation

Maren Schnieder

*Department of Civil and Environmental
Engineering Bochum University of Applied
Sciences*

Bochum, Germany

Maren.Schnieder@hs-bochum.de

<https://orcid.org/0000-0002-9974-2666>

Ana Isabel Canhoto

*Department of Management
University of Sussex Business School
Brighton, UK*

a.i.canhoto@sussex.ac.uk

<https://orcid.org/0000-0002-1623-611X>

Ramin Behbehani

*Brunel Business School
Brunel University London
London, UK*

ramin.behbehani@brunel.ac.uk

<https://orcid.org/0000-0001-8752-5915>

Ahmad Beltagui

*Operations & Information Management
Aston University
Birmingham, UK*

a.beltagui@aston.ac.uk

<https://orcid.org/0000-0003-2687-0971>

Niraj Kumar

*Essex Business School
University of Essex
Southend on Sea, Essex, UK*

n.kumar@essex.ac.uk

<https://orcid.org/0000-0002-7918-5188>

Amirreza Alizamani

*Department of Management
University of Sussex Business School
Brighton, UK*

A.Alizamani@sussex.ac.uk

Abstract— The rapid proliferation and widespread accessibility of electronic word of mouth (eWoM) have sparked a wave of interest in leveraging these sources to guide product and service innovations. While breakthroughs in natural language processing (NLP) are rapidly accelerating, it is important to understand how these data sources themselves differ from more traditional options such as surveys, interviews and focus groups. This paper compares focus groups and online reviews to identify their differences and to highlight how both can complement each other. The study explores both data sources using a thematic analysis and zero-shot text classification, respectively. Results show that prompted (i.e. focus groups) and organic sources (i.e. online reviews) can yield different insights in terms of coverage, granularity, timespan and so forth. Neither one nor the other emerges as a superior data source in this study, as they appear complementary. As online reviews are often written after a delightful or unpleasant experience, their detailed and timely insights may position them as a viable choice for incremental innovation. While focus groups may be the ideal environment for participants to imagine and explore radical innovations, many of the recommendations, as well as concerns, were at least implied in a handful of online reviews. Considering the abundance of online reviews and advances in NLP, even a negligible share may be sufficient to derive insights for radical innovation.

Keywords— *Innovation, product development, customer insight, electronic word of mouth, online reviews, focus groups*

I. INTRODUCTION

A leading cause of innovation failure is poor understanding of customers' needs or preferences [1]. A deeper understanding of consumers enables firms to anticipate and react to external changes [2] and to correctly estimate the market opportunity associated with an innovation project [3]. The first step in developing customer insight is to generate feedback [4].

Evidence shows that information obtained directly from customers is likely to result in successful innovation [5], [6], while the lack of understanding customers' needs is often blamed for a product's failure [1], [7]. Consequently, the development of high-quality customer insight for innovation is a high priority for both academia and practice [8].

There is a growing body of work discussing the advantages and disadvantages of either (i) asking questions directly to consumers in e.g. focus groups or (ii) analysing the feedback that they provide organically in e.g. online reviews or social media posts, to understand customers' attitudes and behaviours, and to support strategic decision making. However, the findings are sometimes contradictory regarding the relative merits of each type of approach, and when one might be preferred over the other. These contradictions and ambiguity mean that it is difficult to develop a theory regarding the situational effectiveness of asking (focus groups) vs listening to (online reviews) opinions, in generating customer insight for innovation research. This is key to making informed decisions regarding which one to use for different types of innovation research. To address these limitations and challenges, this paper provides an additional perspective to the academic literature by comparing the insights generated from electronic word of mouth (eWoM) and focus groups. Specifically, this paper investigates the following research questions:

RQ1: How does prompted customer insight (i.e. focus groups) differ from organic customer insight (i.e. online reviews)?

RQ2: What are the implications of these differences for incremental or radical innovation?

II. RESEARCH BACKGROUND

A. The sources of customer feedback and insight

The first step in developing customer insight is the generation of customer data [8] as this allows product developers to innovate based on factual and meaningful information [4]. A customer's input can help these managers to understand how innovation may be able to solve customers' problems [9]. They also tend to have more ideas about how to improve the product, albeit less radical ones, than those originating from business partners [10]. In addition, customers can provide nuanced information about the context of their consumption [11], which shapes their perception of focal aspects of the solution [12] and is an essential component for successful innovation. For these reasons alone, collecting data directly from customers may improve a competitive advantage [13]. Thus, the acquisition of consumer insight is a key activity in innovation [8], even in the age of artificial intelligence (AI) [14].

Customer insight may be obtained by asking questions directly from consumers, using tools such as focus groups [15]. Alternatively, managers can analyse feedback provided organically by customers, via online review websites [16], among other sources. However, the literature is ambiguous regarding the superiority of one approach over the other. Indeed, some of the benefits and limitations mentioned seem to be common to both approaches. For example, Smithson [17] stated that focus groups allow for the identification of a wide range of views, whereas Kohn [18] argued that they can only provide narrow insight. In turn, Yoo et al. [19] praised online reviews for providing firsthand knowledge, whereas Guo et al. [20] alerted to the existence of fake reviews. Although the consensus appears to be that online reviews are the low-cost alternative to focus groups, Choi et al. [21] claimed that focus groups are a cost-effective source of customer insight. This mirrors, for example, Bigne et al. [22]'s assessment of online reviews as both inexpensive and readily available. Likewise, Kleijnen et al. [23] warned that insight obtained via focus groups may not be representative of the wider customer base, a problem that is also flagged by Wang et al. [24] regarding online reviews.

The contradictions and ambiguity in the existing literature mean that it is a challenging task to develop a theory regarding the situational effectiveness of prompted versus organic feedback for innovation research. Moreover, it is difficult to assess how the insights from each of these approaches might align, diverge or complement each other. Therefore, there is a complexity in making informed decisions regarding which one to use for different types of innovation research.

B. Comparison of online reviews and focus groups

Most academic publications either use traditional data sources, such as focus groups, to guide new product development, or they use online sources such as eWOM or social media posts. Considering the rarity for academics to combine both methods in new product development, it is key to highlight the differences between both sources. A small selection of studies compared the insights generated from online content with those generated from focus groups. Kabel et al. [25]

systematically compared netnography and focus groups. They evaluated 7 focus group transcripts and 26 online discussion boards in the medical domain. Even though they argued that the same customer needs were captured from both data sources, certain aspects tend to be overrepresented in netnography at the expense of other aspects. They ascertain that both are prone to 'group think' and group dynamics; the evidence being even stronger in online communities. While netnography captures current requirements, focus groups concentrate on historical usage. They emphasised the need for organisations to employ multiple research methods, instead of solely relying on one.

Focusing on usability assessment, Cowley et al. [26] also compared two focus group sessions with two online forum groups. While more questions were answered in the online forums, the number of suggested design improvements was comparable. Focus group responses provided more contextual information, sometimes even going off topic. The responses to the online forum tend to be on point. They join Kabel et al. [25], in highlighting the differences between both data sources.

III. RESEARCH DESIGN

A. Data collection: Focus groups

Ten mini-focus group sessions were carried out between December 2023 and January 2024; each lasting approximately one hour. Conducting the focus groups synchronously online enhances scheduling flexibility and expands the pool of participation. Only 2 to 3 participants joined each group to maintain the participants' engagement in the online discussions. These were selected using non-probability, purposive sampling, to recruit a wide range of consumers.

A set of semi-structured questions and descriptions of the business models were defined and refined through 4 pilot focus groups, including one in-person session, to ascertain whether the nature of the participants' interpersonal dynamics in the online groups was similar to the one in-person. In the group interviews, the moderator (i) introduced the clothes cleaning solutions one at a time, (ii) checked whether the participants had experience with the solution, and (iii) answered any clarification questions the participants might have, before the participants discussed the advantages as well as disadvantages and formulated suggestions for improvement.

The group interviews were recorded, and the transcripts were anonymised and cleaned manually by rewatching the recordings.

B. Data collection: Online reviews

As an example of organic feedback, over 2000 online reviews published on Trustpilot and Google Maps were used. Trustpilot is a well-known third-party review website [27], launched in the UK in 2014 [28]. Being designed as an open platform, it encourages customers to write a review about their recent experiences with buying or accessing services from a company [28]. While companies can invite customers to write reviews on Trustpilot, they can also do so unprompted. The reviews comprise a Likert Scale rating of the company, from 1 star (i.e. bad) to 5 stars (i.e. excellent), as well as a comment describing the experience with the company [27]. Trustpilot has commonly been used as a source for reviews in the academic

literature, from pet-food subscriptions [29] and second-hand fashion market [27] to multilevel marketing [30].

The reviews of major UK online retailers selling washing machines, as well as operators of launderettes, were sourced from Trustpilot (<https://uk.trustpilot.com>, accessed: 10.05.2024 to 03.06.2024). In the case of online retailers, the reviews were limited to those mentioning washing machines.

Additional reviews on launderettes were sourced from Google Maps (<https://www.google.co.uk/maps/>, accessed: 31.05.2024). Such reviews have been used to assess passenger satisfaction with transport hubs (e.g. airports/train stations [31], [32]), medical services (e.g. vaccination centres [33], emergency departments [34]), as well as other locations such as parks [35], co-working spaces [36] and electric vehicle fast charging stations [37].

C. Data analysis

Given that the methods usually differ to examine online reviews and focus groups, both were initially analysed using their usual method before reversing them. By doing so, qualitative and quantitative approaches were combined to conduct a comparative analysis, as recommended by Ragin [38]. A few emerging studies have also highlighted the benefits of combining qualitative insights gained from human coding with quantitative insights from large-scale data-driven natural language processing techniques to ‘read consumer’s minds’ (e.g. [39]). For the qualitative analysis, the focus group data were analysed using NVivo 14 applying a systematic three-phase framework designed to support abductive reasoning [40], while combining aspects of thematic and qualitative content analysis methodologies. A deductive content analysis was performed first [41]. Instances that did not match any of the theory-derived codes were coded inductively, as a new category or a sub-category of an existing code [42]. Secondly, the thematic analysis approach outlined by Fereday et al. [43] was applied. This phase harmonised the emerging and theory-driven codes. Finally, themes were synthesized and their descriptions formulated.

To verify the prediction accuracy of the zero-shot text classification, 377 consumer reviews were manually labelled. The labelling was not guided by the themes identified in the focus groups. Instead, an inductive approach was chosen, allowing the codes to emerge instead of fitting the reviews into a pre-defined framework. Not every aspect mentioned was labelled as the focus was on those relevant for product innovation. Only aspects included at least 26 times were considered to ensure that sufficient labelled reviews were available to verify the reliability of the classification. The Zero-shot text classification was implemented in PyTorch using HuggingFace. A vast array of candidate labels was tested using facebook/bart-large-mnli, a commonly used transformer-based model introduced by Lewis et al. [44]. Employing a method known as Natural Language Inference (NLI), the model is prompted with a premise (i.e. a sentence of the review) and a hypothesis (i.e. a candidate label). The model predicts the probability that the hypothesis aligns with the premise [45]. The hypothesis used was modified as follows: “the review mentions {}”. One sentence was classified at a time, and the highest probability amongst all sentences of a review was used for each

label and review, respectively. Based on the set of candidates for each label, the most suitable one for each aspect was selected based on the f1-score. The threshold was optimised for each candidate label to maximise the f1-score, similar to Van Nooten et al. [46]. Next, various models, i.e. MoritzLaurer/DeBERTa-v3-base-mnli-fever-anli, MoritzLaurer_mDeBERTa-v3-base-mnli-xnli, and MoritzLaurer/DeBERTa-v3-large-mnli-fever-anli-ling-wanli, introduced by Laurer et al. [47], have been compared. Only the best candidate label for each aspect was used if the f1-score was above 0.8. Up to four candidate labels that had an f1-score of above 0.5 were used for the other aspects. For low performing aspects, it could not be said with certainty which candidate label would perform well enough. Finally, the better performing model with the most suitable set of labels was used to classify all reviews (i.e. also those that were not labelled).

The same process was repeated focusing on laundrettes where 209 reviews were manually labelled (around half from each data source). All labels that were mentioned at least 15 times were included.

The labels and the best-performing model were then applied to the focus group transcripts. Since no test data was available, a suitable threshold for each label was determined manually. A combination of zero-shot text classification using facebook/bart-large-mnli and manual annotation was used to identify whether suggestions for improvements, or concerns raised in the focus groups, were also mentioned in at least one online review.

D. Ethical considerations and limitations

Many studies analysing online content state that no ethics application was deemed necessary, or the study was reviewed but considered exempt, given that there is no reasonable expectation of privacy [48] [49] [50] [51]. Following this, no ethics application was submitted.

Reviews of online retailers selling a product might be less comparable to focus groups discussing that product. However, this comparison was still included due to the vast availability of such online reviews and the resulting interest in utilising these. Online reviews and focus group discussions on launderettes are readily more comparable in this respect. However, some might argue that the differing insights might stem from those writing online reviews being more frequent users of laundrettes than most of the focus group participants in this study. Considering the word choice in some of the negative online reviews, or the blatant statements, it appears that some are using laundrettes due to a lack of choice in alternative options (e.g. “AVOID. I shower with my clothes on to avoid having to deal with your service! [...]” or “[...] I have now resorted to washing my clothes in my sink [...]).

IV. RESULTS

A. Participants’ profiles: Online and focus group

The focus groups brought together 29 participants (59% female), of whom 38% lived with children, 24% with a partner, 21% lived with other adults and 17% alone. The participants’ age was spread evenly (18-25: 10%, 26-35: 21%, 36-46: 31%, 46-65: 38%), the majority had a washing machine at home (72%) and were employed (68%).

Tracking customers across online platforms or accessing their online profiles was not possible in this study. Hence, demographic information related to the online reviews' dataset could only be inferred indirectly from the posts themselves [52]. It was possible to identify some contextual information (e.g. "elderly": "I ordered [a] washing machine on line for my elderly mother who lives on her own"; employed: "Compensation for [...] [my] wasted day of holiday was an insulting £10 voucher [...]"; living in a flat: "They happily took [the washing machine] up stairs to my flat."; being a student: "they know full well [...] [their] customers are students with no money"). However, the availability of such information was insufficient to develop a comprehensive understanding of the characteristics of the customers.

B. Zero-shot text classification: Washing machine

The performance of each classifier – applied to washing machine reviews - is illustrated in Table I.

TABLE I. PERFORMANCE OF VARIOUS ZERO-SHOT TEXT CLASSIFIERS

<i>Model</i>	<i>Hamming Loss</i>	<i>f1 micro</i>	<i>f1 macro</i>	<i>Precision micro</i>	<i>Recall micro</i>
facebook/bart-large-mnli	0.076	0.853	0.800	0.839	0.867
MoritzLaurer/DeBERTa-v3-base-mnli-fever-anli	0.103	0.809	0.749	0.765	0.858
MoritzLaurer_mDeBERTa-v3-base-mnli-xnli	0.119	0.776	0.696	0.744	0.810
MoritzLaurer/DeBERTa-v3-large-mnli-fever-anli-ling-wanli	0.062	0.881	0.854	0.857	0.905

The per-label performance of MoritzLaurer/DeBERTa-v3-large-mnli-fever-anli-ling-wanli is illustrated in Table II. The price was unusually difficult to predict, given that reviews focused on refunds were also classified as talking about money by the zero-shot text classification, whereas the authors split refund and price/affordability into two different categories. The label for the removal of the washing machine is excessively long. Any shorter version led to the model predicting the removal of packaging incorrectly as positive - an issue still partially present with the currently used label.

TABLE II. PER LABEL PERFORMANCE OF MORITZLAURER/DEBERTA-V3-LARGE-MNLI-FEVER-ANLI-LING-WANLI (WASHING MACHINE)

<i>Label</i>	<i>Precision</i>	<i>Recall</i>	<i>f1</i>
Pay	0.922	0.849	0.884
Delivery experience	0.923	0.983	0.952
Upstairs	0.900	0.818	0.857
Money back	0.921	0.972	0.946
Installation or fitting	0.797	0.944	0.864
Water leaking problem	0.750	0.900	0.818
Recycling or removing or disconnecting of washing machine	0.689	0.697	0.693
In stock or out of stock	0.867	0.813	0.839
Insurance or warranty or guarantee	0.756	0.816	0.785
Packaging	0.854	0.953	0.901

Table III shows the description of each label and its share in the entire dataset. It should be noted that the share of reviews mentioning a specific aspect does not necessarily indicate its importance [37]. As reviews are commonly limited to what is on a person's mind, it may be unlikely that they mention something that could have been a problem if it was not. For example, EV drivers might not point out a lack of lighting at EV charging stations if they charged during the day in this instance [37]. People may not mention something they appreciate if they are unaware that this could be on offer. For example, a park at EV charging stations was a reason for praise when available, while no one commented on this when it was not present [37]. In short, the share of reviews mentioning an aspect does not necessarily relate to its overall importance.

TABLE III. OVERVIEW OF ASPECTS MENTIONED IN THE REVIEWS (WASHING MACHINE)

<i>Label</i>	<i>Meaning</i>	<i>Share</i>
Delivery experience	Delivery of the washing machine	74%
Installation or fitting	Installation of the washing machine	43%
Pay	Affordability and price paid	32%
Recycling or removing or disconnecting of washing machine	Removal of the old washing machine	19%
Money back	Requests for refunds	13%
Packaging	Packaging of the washing machine	9%
Water leaking problem	Mentions of a water leak or dripping	8%
Insurance or warranty or guarantee	Warranty or insurance	9%
In stock or out of stock	Stock availability	8%
Upstairs	Challenges of delivering the washing machine upstairs	3%

The content covered per person was higher in the focus groups, most likely due to the much longer transcripts compared to online reviews. Unsurprisingly, all participants touched upon the price as well as the installation of washing machines. For those writing an online review, the delivery experience was the most covered topic. Only two-thirds of the focus group participants remembered challenges with the delivery process. Slightly more demanded a warranty, while most of those writing an online review luckily did not have to think about the warranty so shortly after their purchase. Water leaks, or the removal of a washing machine, were on the mind of around half of the participants, while significantly fewer mentioned these in the online reviews.

C. Zero-shot text classification: Launderette

Since the per-label performance varied greatly sometimes, both facebook/bart-large-mnli and MoritzLaurer/DeBERTa-v3-large-mnli-fever-anli-ling-wanli were utilised (Table IV).

TABLE IV. PER LABEL PERFORMANCE (LAUNDERETTE)

<i>Label</i>	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>f1</i>
Apps ^a	0.970	0.970	0.970	0.970
Refund ^a	0.968	0.938	1.000	0.968
Price ^a	0.891	0.869	0.915	0.891
Broken ^a	0.841	0.866	0.817	0.841
Staff or customer service ^a	0.838	0.727	0.989	0.838
Fraud ^a	0.867	0.788	0.963	0.867
Drying ^a	0.855	0.756	0.985	0.855
Clean or tidy ^b	0.779	0.682	0.909	0.779

<i>Label</i>	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>f1</i>
Large or small size ^b	0.667	0.765	0.591	0.667
Detergent or washing powder ^b	0.800	0.800	0.800	0.800
Coins or change ^b	0.757	0.683	0.848	0.757
Parking space or parking spot ^b	1.000	1.000	1.000	1.000
Card or contactless payment ^b	0.769	0.909	0.667	0.769

^a MoritzLaurer/DeBERTa-v3-large-mnli-fever-anli-ling-wanli

^b facebook/bart-large-mnli

TABLE V. OVERVIEW OF ASPECTS MENTIONED IN THE REVIEWS (LAUNDRETTES)

<i>Label</i>	<i>Meaning</i>	<i>Share</i>
Apps	Apps	15%
Refund	Refunds	22%
Price	Price	69%
Broken ^a	Broken washing machines, not functioning apps, card machines, tumble driers etc.	37%
Staff or customer service	Staff or customer service	68%
Fraud	Comments like daylight robbery or scam	18%
Drying	Drying and tumble dryers	49%
Clean or tidy	Cleanliness of the launderette, the washing machines, or clothes after being washed	17%
Large or small size	Size or capacity of washing machines	8%
Detergent or washing powder	Availability of detergents	10%
Coins or change	Coins and cash	25%
Parking space or parking spot	Parking spots	5%
Card or contactless payment	Cards and contactless payments	10%

^a Also mentions of washing machines broken at home were classed as positive.

D. Thematic analysis: Online reviews and focus groups

To refrain from publishing the same results multiple times, the reader is referred to other publications, that are part of this project, for a detailed overview of the results of the focus groups (e.g. [53]).

The participants mentioned a vast array of worries about laundrettes (e.g. clothes being stolen, or the safety of the neighbourhood) in the focus groups. Their suggestions for improvements ranged from cafés to entire arcades. In terms of washing machines, the participants mentioned aspects such as energy/water usage, eco-cycle, or longevity. While these aspects are sometimes rather rarely mentioned in online reviews, it was possible for most of these to identify at least one review that hinted at it. Some examples of these are listed below:

a) *Interior design*: “[...] They even have lovely quirky decor! [...]”

b) *Wi-Fi*: “[...] could use a little refurbishment soon or maybe free Wi-Fi for customer”

c) *Stolen*: “[...] my clothes all messed up. [...] Still not sure if any of my clothes were stolen. [...]”

d) *Entertainment*: “[...] The area is interesting with crafty shops and café [...]”

e) *Washing dog/horse related items*: “[...] They are perfect for cleaning towels after a long muddy dog walk and general freshening up of my smelly dog’s bed covers [...]”

f) *Neighbourhood*: “[...] Interesting street to explore. Excellent café down the street”

g) *Safety*: “[...] Also the area is a safe place [...] full [...] [with] nice people and places to eat [...]”

h) *Cameras*: “[...] CCTV so feel safe even when I’ve used the machines at night. [...]”

i) *Lost socks*: “[...] socks get stuck in the edge which is mouldy, [...]”

j) *Waiting area*: “[...] There is a pleasant waiting area [...]”

k) *Folding area*: “[...] a nice folding table at the rear [...]”

l) *Clothes being ruined*: “[...] ruined the clothes that were in the wash (they came out shrunken and the color washed out) [...]”

m) *time commitment*: “said it will take 36 minutes I then spent over 2 hours waiting for it to wash total joke.”

n) *Instructions*: “[...] the whole thing has unclear instructions and is totally not user-friendly. [...]”

There are rare exceptions where it was not possible to identify a similar reference to an aspect in the online reviews. One example is privacy concerns (i.e. demand for cubicles so that their underwear would not be seen by others). However, one person complained in the online reviews that the launderette operator “Request unreasonable amounts of personable [personal] data”. Not having to worry about the electricity bill when using launderettes was mentioned a few times in the focus groups and somewhat acknowledged in the online reviews: “[...] Making it a lifesaving shop [...] [if] you haven’t got enough electric at home, [...]”. Other advantages that were occasionally mentioned were the ability to save space at home, as well as environmental benefits. While it was not possible to identify the first advantage in the online reviews, the latter is only mentioned if it coincides with financial benefits “[...] I’ve resorted to using the eco wash, which is a measly 50p cheaper, [...]”.

E. Timeline covered and level of detail: Washing machines

The focus group discussions covered the entire lifecycle of owning a washing machine, from purchase decision to the use phase, including maintenance and eventual disposal. It was common to find references to multiple washing machines owned during someone’s lifetime. In contrast, the online reviews focused on a comparatively small time-frame – usually the period just before writing the review. This is most likely due to them often being written after a recent purchase; sometimes prompted by the seller, other times by a particularly delightful or unpleasant experience. The focus is usually limited to the delivery and installation process, as this is often the most recent interaction with the seller. Some also included (i) details of the ordering process beforehand, (ii) the first few washes, or (iii) dealing with the customer support when things go wrong. While a few wrote the reviews years after the purchase, or compared their recent purchase with previous experiences, others had not even received the delivery yet, before they felt compelled to write a review in disappointment. In terms of removal, some mentioned the pick-up of their old washing machine, with one stating that they “[...] ordered a washing machine [...] and taking away [of the] old machine [...]”.

Another distinctive feature of online reviews is the sporadic provision of rather detailed information, such as a specific

washing machine (e.g. “[...] Bosh series 4 washing machine [...]”) or names of the staff they interacted with. Depending on the source of the online reviews, it is possible to even know the details of the exact purchase (e.g. date, product etc.) related to the comment (e.g. eBay). Unless prompted by the moderator, such detailed information might not be volunteered – or even remembered – by focus group participants. At most, brand names were only mentioned in the focus groups, but certainly no specific details apart from the capacity, with one stating “[...] it's only a [...] five KG [...]”. However, it might not be possible to expect consumers to remember the model number of the washing machine. The detailed purchase information provided in some online reviews can be extremely helpful to pinpoint “the weakest link” and identify trends.

In summary, the narrow time focus of online reviews offers insight into specific aspects of the most recent experience, with a focus on unfulfilled expectations and staff interactions, while at other times providing very specific details. In turn, the focus groups support the identification of important aspects across the customer journey, and sometimes across multiple washing machines, while providing general and reflective feedback. Reviews of online retailers can offer insight into ancillary aspects of the purchase, rather than the machine itself. This kind of insight can help the manager understand how ancillary aspects can offer moments of delight or, more often, be a source of frustration. In turn, the focus groups support the identification of important aspects across the life cycle of owning a washing machine, or multiple units, providing a reflective view. This indicates that neither data source is an ideal substitute for the other, as far as customer insight is concerned.

F. Timeline covered and details: Launderettes

The reasons given for the (previous/current) use of a launderette were similar across online reviews and focus groups, including permanent ones such as lack of a washing machine at home, and temporary ones such as the need to wash large bedding or being on holiday.

All the focus group comments referred to several uses of a launderette in the present or past. Online reviews were divided between reports on specific instances (i.e. first/recent/most memorable visit), and recurring experiences. For example, one review stated that “Every time I go to use the washing machine it's never working due to the card payment system”. While another one stated “[...] I've probably visited 100+ self-service laundrettes throughout Europe over the years, this is the first time I've ever had a problem. [...]”.

The level of detail provided was slightly different. Given that the times of using launderettes were for many focus group participants, a few years ago, the feedback tended to be more reflective on the entire experience, as opposed to pinpointing a specific instance. One participant, for example, remembered that “[...] you have to walk all the way to the laundrette, carry all your stuff. That's a bit sort of inconvenient, and I suppose maybe [...] [the] lack of privacy because you've got your underwear. [...]”. Others mention their worries in general, even if they may never have happened in reality “[...] I also don't know the previous users. You don't know how dirty their washing might be, and even though it is a washing machine, if the last wash was really dirty, it might affect my wash or if they wash dark clothes

in the previous wash and I wash a light wash [...] all these little uncertainties that will affect my own wash. [...]” While specific events were sometimes mentioned for an illustrative purpose in the focus groups, the detailed description, transcript-like, of specific situations in launderettes are certainly more likely to be seen in online reviews: “[...] I used for the first time today. I was washing [...] bedcovers. I set my wash which was 37 minutes. I returned 5 mins early and the wash cycle had ended. The wash was not clean and had not been spun dry. [...]”.

Another important difference was the information that could be derived to improve laundrettes. Participant suggestions included cubicles for privacy, arcades for entertainment, and cafés for relaxing or working. Their worries ranged from their personal safety to their clothes being stolen or ruined. These recommendations and concerns are only present in a negligible number of reviews, as most are focused on improving the current service instead of exploring additional offers.

In short, the timeframe covered in online reviews of the launderettes is varied. A few have a similar scope to focus groups. However, many were written in the heat of the moment after a particular instance (e.g. negative, positive or their first visit), and in these cases, often provide specific details. Some online reviews also focus on only one distinct theme. Focus groups provide a more reflective summary of their experiences; sometimes, these go beyond the service itself. Focus groups also offer an insight into the worries consumers may have that prevent them from using the service and allow them to explore options to overcome these barriers. Focus groups can be used to re-imagine launderettes, while online reviews may primarily help the management to understand the current issues experienced.

G. Role taking/Altruistic perspective

The focus group participants were able to “remember the old days” as well as put themselves into the shoes of someone who currently uses a launderette (e.g. their children), or imagining situations where they would be happy to use it (e.g. “[...] if it's a closed community, I would probably worry less in terms of the state of the machines, the hygiene and cleanliness, sanitation of the machines, and things like that.[...]”). As those writing online reviews appear to be a current or recent customer, they usually focus on their own experience in the reviews. One exception is those writing online reviews on behalf of someone else, like their elderly relative or child going to university.

While participants obviously explored their own service outcomes in terms of buying a washing machine, many also mentioned the perspective of others. They also discussed the suitability of the product if they were at another stage in their life, or economic situation. A few participants argued that they are buying new washing machines simply because that is what their parents did. While there are exceptions in the online reviews, consumers usually focus on their own experiences.

In short, even though online reviews are usually centred around the experience of the individual person, this is not a major issue due to the abundance of them. Participants' ability “to put on many hats” allows for the generation of a rich data source of experiences, despite the usually small number of participants. Another advantage of focus groups is the

opportunity to receive feedback on situations where the product and service would not be suitable; or explore how to overcome these barriers. Such detail is difficult to extract from online reviews, given that those for whom the service is unsuitable will unlikely remain, or never be, a customer.

V. DISCUSSION AND RELATION TO PRIOR WORK

The conclusions drawn in this study align with those of Kabel et al. [25], who argued that netnography may provide rich data sources; albeit some information may be over- or underrepresented compared to focus groups. As prompted by the moderator, the focus groups covered the entire customer journey and the life cycle of multiple products or services. In contrast, online reviews were often written in the heat of the moment, capturing what is in the mind of the consumer at that time. A conclusion reflected in Kabel et al. [25], who argued that online discussion boards may overrepresent occasional but extraordinary situations. However, Kabel et al. [25] concluded that online discussion boards prioritise more details of their personal lives and emotional topics than focus groups. This was not evident in the online reviews included in this study. While a few might provide a detailed backstory in their reviews, essentially everyone in the focus groups was talking about their personal experiences and those of those they know. However, this might be caused by the difference in online data source used (i.e. online reviews (this study) vs. online discussion boards [25]).

VI. CONCLUSIONS

This study compares the insights derived from online reviews (i.e. organic) and focus groups (i.e. prompted). Prompted and organic sources of customer insight can provide significantly different information in terms of, for example, scope, detail, and covered timeframe.

In general, the results of this study indicate that online reviews may be more useful for incremental innovation of mature products, for which there is extensive consumer experience to derive insight from. Considering the vast availability of online reviews and the detailed granularity, they can be a source of information related to pain points only experienced by a handful of customers. In the case of novel offerings and radical innovation, the focus groups may offer an ideal environment to imagine, and an opportunity to elaborate on their concerns and suggestions.

Neither one nor the other emerges as a superior data source in this study, as they appear complementary. Hence, this study provides further evidence for Kabel et al. [25]'s statement recommending companies to rely on multiple data sources. While the advances of AI will certainly improve data analysis methods, differences in the data sources may not be easily overcome. However, many of the recommendations, as well as concerns raised in focus groups, were at least implied in one or a handful of online reviews. With the plentiful availability of online reviews and advances in NLP, even the smallest share may be sufficient to extract useful insights for radical innovation.

ACKNOWLEDGMENT

This work was supported by the UKRI Made Smarter Innovation Challenge and the Economic and Social Research Council via InterAct [Grant Reference ES/W007231/1].

REFERENCES

- [1] D. Baxter, P. Trott, and P. Ellwood, "Reconceptualising innovation failure," *Res Policy*, vol. 52, no. 7, Sep. 2023, doi: 10.1016/j.respol.2023.104811.
- [2] J. P. Deschamps, "Classic root causes of innovation failures-things we all know but sometimes forget," in *Strategy and Communication for Innovation: Integrative Perspectives on Innovation in the Digital Economy*, Springer International Publishing, 2017, pp. 41–60. doi: 10.1007/978-3-319-49542-2_4.
- [3] I. Jenson, P. Leith, R. Doyle, J. West, and M. P. Miles, "Innovation system problems: Causal configurations of innovation failure," *J Bus Res*, vol. 69, no. 11, pp. 5408–5412, Nov. 2016, doi: 10.1016/j.jbusres.2016.04.146.
- [4] M. Wedel and P. K. Kannan, "Marketing analytics for data-rich environments," *J Mark*, vol. 80, no. 6, pp. 97–121, Nov. 2016, doi: 10.1509/jm.15.0413.
- [5] K. E. Gruner and C. Homburg, "Does Customer Interaction Enhance New Product Success?," *J Bus Res*, vol. 49, no. 1, 2000.
- [6] A. S. Cui and F. Wu, "Utilizing customer knowledge in innovation: antecedents and impact of customer involvement on new product performance," *J Acad Mark Sci*, vol. 44, no. 4, pp. 516–538, Jul. 2016, doi: 10.1007/s11747-015-0433-x.
- [7] M. Falahat, S. C. Chong, and C. Liew, "Navigating new product development: Uncovering factors and overcoming challenges for success," *Heliyon*, vol. 10, no. 1, 2024.
- [8] S. Stremersch, E. Cabooter, I. A. Guitart, and N. Camacho, "Customer insights for innovation: A framework and research agenda for marketing," *J Acad Mark Sci*, vol. 53, no. 1, 2025, doi: 10.1007/s11747-024-01051-8.
- [9] R. Adams, S. Quinton, and R. Pera, "Social media representations of innovation by non-users: everyday problem solving," in *The ISPIM Innovation Conference – Innovating Our Common Future*, Berlin, Jun. 2021.
- [10] A. Ordanini and A. Parasuraman, "Service innovation viewed through a service-dominant logic lens: A conceptual framework and empirical analysis," *J Serv Res*, vol. 14, no. 1, pp. 3–23, Jan. 2011, doi: 10.1177/1094670510385332.
- [11] D. L. Roberts, F. T. Pillar, and D. Lüttgens, "Mapping the Impact of Social Media for Innovation: The Role of Social Media in Explaining Innovation Performance in the PDMA Comparative Performance Assessment Study," *Journal of Product Innovation Management*, vol. 33, pp. 117–135, Dec. 2016, doi: 10.1111/jpim.12341.
- [12] K. B. Yap and J. C. Sweeney, "Zone-of-tolerance moderates the service quality-outcome relationship," *Journal of Services Marketing*, vol. 21, no. 2, pp. 137–148, 2007, doi: 10.1108/08876040710737895.
- [13] E. J. Hultink, K. Talke, A. Griffin, and E. Veldhuijzen, "Market information processing in new product development: The importance of process interdependency and data quality," *IEEE Trans Eng Manag*, vol. 58, no. 2, pp. 199–211, May 2011, doi: 10.1109/TEM.2009.2034254.
- [14] R. G. Cooper, "SMEs' use of AI for new product development: Adoption rates by application and readiness-to-adopt," *Industrial Marketing Management*, vol. 126, pp. 159–167, Apr. 2025, doi: 10.1016/j.indmarman.2025.01.016.
- [15] A. Garrett and C. Wrigley, "Navigating market opportunity: traditional market research and deep customer insight methods," *Qualitative Market Research*, vol. 22, no. 3, pp. 456–471, Jun. 2019, doi: 10.1108/QMR-05-2017-0091.
- [16] K. Lu, T. Zou, and J. Du, "Two forms of customer involvement and new product development performance in the digital context: The moderating role of new product development stage," *Technovation*, vol. 134, Jun. 2024, doi: 10.1016/j.technovation.2024.103023.
- [17] J. Smithson, "Using and analysing focus groups: limitations and possibilities," *Int J Soc Res Methodol*, vol. 3, no. 2, p. 19, 2000, [Online]. Available: <http://www.tandf.co.uk/journals>
- [18] K. Kohn, "Idea generation in new product development through business environmental scanning: The case of XCar," *Marketing Intelligence and*

- Planning*, vol. 23, no. 7, pp. 688–704, 2005, doi: 10.1108/02634500510630212.
- [19] K. H. Yoo and U. Gretzel, “What motivates consumers to write online travel reviews? What motivates consumers to write online travel reviews? Recommended Citation Recommended Citation,” *Information Technology & Tourism*, vol. 10, no. 4, pp. 283–295, 2008, [Online]. Available: <https://ro.uow.edu.au/commpapers/1540>
 - [20] J. Guo, X. Wang, and Y. Wu, “Positive emotion bias: Role of emotional content from online customer reviews in purchase decisions,” *Journal of Retailing and Consumer Services*, vol. 52, Jan. 2020, doi: 10.1016/j.jretconser.2019.101891.
 - [21] J. J. Choi, C. A. Robb, M. Mifli, and Z. Zainuddin, “University students’ perception to online class delivery methods during the COVID-19 pandemic: A focus on hospitality education in Korea and Malaysia,” *J Hosp Leis Sport Tour Educ*, vol. 29, Nov. 2021, doi: 10.1016/j.jhlste.2021.100336.
 - [22] E. Bigne, M. Zanfardini, and L. Andreu, “Destination online reviews: lexicons and thematic arguments on social responsibility,” *Tourism Review*, vol. 80, no. 4, pp. 944–965, Mar. 2025, doi: 10.1108/TR-10-2023-0725.
 - [23] M. Kleijnen, N. Lee, and M. Wetzels, “An exploration of consumer resistance to innovation and its antecedents,” *J Econ Psychol*, vol. 30, no. 3, pp. 344–357, Jun. 2009, doi: 10.1016/j.joep.2009.02.004.
 - [24] W. Wang, Y. Feng, and W. Dai, “Topic analysis of online reviews for two competitive products using latent Dirichlet allocation,” *Electron Commer Res Appl*, vol. 29, pp. 142–156, May 2018, doi: 10.1016/j.elerap.2018.04.003.
 - [25] D. Kabel, J. Martin, M. Elg, and L. Witell, “Capturing the voice of the customer: focus groups versus netnography?,” *Total Quality Management and Business Excellence*, vol. 35, no. 11–12, pp. 1359–1377, 2024, doi: 10.1080/14783363.2024.2369592.
 - [26] J. A. Cowley and J. Radford-Davenport, “Qualitative data differences between a focus group and online forum hosting a usability design review: A case study,” in *Proceedings of the Human Factors and Ergonomics Society*, 2011, pp. 1356–1360. doi: 10.1177/1071181311551282.
 - [27] H. M. R. ul Hasan, C. Lang, and S. Xia, “Investigating Consumer Values of Secondhand Fashion Consumption in the Mass Market vs. Luxury Market: A Text-Mining Approach,” *Sustainability (Switzerland)*, vol. 15, no. 1, Jan. 2023, doi: 10.3390/su15010254.
 - [28] S. Littlechild, “Exploring customer satisfaction in Great Britain’s retail energy sector part III: A proposed Overall Customer Satisfaction score,” *Util Policy*, vol. 73, Dec. 2021, doi: 10.1016/j.jup.2021.101299.
 - [29] D. Lima, R. F. Ramos, and P. M. Oliveira, “Customer satisfaction in the pet food subscription-based online services,” *Electronic Commerce Research*, vol. 24, no. 2, pp. 745–769, Jun. 2024, doi: 10.1007/s10660-024-09807-8.
 - [30] H. Saadiah Rusli and N. Farizah Ibrahim, “Exploring the Sentiment and Online Review of Multilevel Marketing (MLM) Company Products,” in *Proceedings of the 11th Annual International Conference on Industrial Engineering and Operations Management*, Singapore, Jul. 2021.
 - [31] M. Schnieder and M. Dora, “Trains or Planes? An Aspect-Based Sentiment Analysis of Travel Satisfaction,” in *2024 IEEE International Conference on Smart Mobility, SM 2024*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 242–247. doi: 10.1109/SM63044.2024.10733556.
 - [32] K. Lee and C. Yu, “Assessment of airport service quality: A complementary approach to measure perceived service quality based on Google reviews,” *J Air Transp Manag*, vol. 71, pp. 28–44, Aug. 2018, doi: 10.1016/j.jairtraman.2018.05.004.
 - [33] S. Danek, M. Büttner, J. Krois, and F. Schwendicke, “How Do Users Respond to Mass Vaccination Centers? A Cross-Sectional Study Using Natural Language Processing on Online Reviews to Explore User Experience and Satisfaction with COVID-19 Vaccination Centers,” *Vaccines (Basel)*, vol. 11, no. 1, Jan. 2023, doi: 10.3390/vaccines11010144.
 - [34] N. Derdzakyan, A. Pourmand, R. Shesser, S. Ganguli, and J. Trevino, “The potential use of Google reviews to assess patient satisfaction in the emergency department,” *American Journal of Emergency Medicine*, vol. 52, pp. 110–113, Feb. 2022, doi: 10.1016/j.ajem.2021.12.013.
 - [35] Y. F. Koh, H. H. Loc, and E. Park, “Towards a ‘City in Nature’: Evaluating the Cultural Ecosystem Services Approach Using Online Public Participation GIS to Support Urban Green Space Management,” *Sustainability (Switzerland)*, vol. 14, no. 3, Feb. 2022, doi: 10.3390/su14031499.
 - [36] M. Schnieder and A. Beltagui, “Understanding Co-Working Space Users: Aspect-Based Sentiment Analysis,” in *9th IEEE/ACIS International Conference on Big Data, Cloud Computing, and Data Science, BCD 2024*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 9–14. doi: 10.1109/BCD61269.2024.10743079.
 - [37] M. Schnieder, “Fast Charging Stations in the USA: A Multi-Label Text Classification of Online Reviews,” in *Proceedings - 2025 Artificial Intelligence for Business, AIx B 2025 (accepted)*, 2025.
 - [38] C. C. Ragin, *The comparative method: Moving beyond qualitative and quantitative strategies*. Univ of California Press, 2014.
 - [39] W. Kwon, “Reading customers’ minds through textual big data: Challenges, practical guidelines, and proposals,” *Int J Hosp Manag*, vol. 111, May 2023, doi: 10.1016/j.ijhm.2023.103473.
 - [40] G. A. Bowen, “Document analysis as a qualitative research method,” *Qualitative Research Journal*, vol. 9, no. 2, pp. 27–40, 2009, doi: 10.3316/QRJ0902027.
 - [41] W. J. Potter and D. Levine-Donnerstein, “Rethinking validity and reliability in content analysis,” *Journal of Applied Communication Research*, vol. 27, no. 3, pp. 258–284, 1999, doi: 10.1080/00909889909365539.
 - [42] H. F. Hsieh and S. E. Shannon, “Three approaches to qualitative content analysis,” *Qual Health Res*, vol. 15, no. 9, pp. 1277–1288, Nov. 2005, doi: 10.1177/1049732305276687.
 - [43] J. Fereday, N. Adelaide, S. Australia, and A. Eimear Muir-Cochrane, “Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development,” *Int J Qual Methods*, vol. 5, 2006.
 - [44] M. Lewis *et al.*, “BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension,” Oct. 2019, [Online]. Available: <http://arxiv.org/abs/1910.13461>
 - [45] M. J. J. Bucher and M. Martini, “Fine-Tuned ‘Small’ LLMs (Still) Significantly Outperform Zero-Shot Generative AI Models in Text Classification,” Aug. 2024, [Online]. Available: <http://arxiv.org/abs/2406.08660>
 - [46] J. Van Nooten, A. Kosar, G. De Pauw, and W. Daelemans, “One Size Does Not Fit All: Exploring Variable Thresholds for Distance-Based Multi-Label Text Classification,” Oct. 2025, [Online]. Available: <http://arxiv.org/abs/2510.11160>
 - [47] M. Laurer, W. Van Attevelde, A. Casas, and K. Welbers, “Less Annotating, More Classifying-Addressing the Data Scarcity Issue of Supervised Machine Learning with Deep Transfer Learning and BERT-NLI,” *Political Analysis*, vol. 32, no. 1, pp. 84–100, 2024, doi: 10.1017/pan.2023.20.
 - [48] V. S. Anoop and S. Sreelakshmi, “Public discourse and sentiment during Mpox outbreak: an analysis using natural language processing,” *Public Health*, vol. 218, pp. 114–120, May 2023, doi: 10.1016/j.puhe.2023.02.018.
 - [49] M. Y. Hameed, L. Al-Hindi, S. Ali, H. K. Jensen, and C. C. Shoults, “Broadening the Understanding of Medical Students’ Discussion of Radiology Online: A Social Listening Study of Reddit,” *Curr Probl Diagn Radiol*, vol. 52, no. 5, pp. 377–382, Sep. 2023, doi: 10.1067/j.cpradiol.2023.04.003.
 - [50] E. L. Eschliman *et al.*, “First-hand accounts of structural stigma toward people who use opioids on Reddit,” *Soc Sci Med*, vol. 347, Apr. 2024, doi: 10.1016/j.socscimed.2024.116772.
 - [51] M. J. Meade, P. Sooriakumaran, X. Ju, D. Hunter, and L. Jamieson, “Evaluation of orthodontic retention and retainer content on the Reddit social media website,” *J World Fed Orthod*, vol. 12, no. 5, pp. 213–219, Oct. 2023, doi: 10.1016/j.ejwf.2023.06.003.
 - [52] J. Li and X. Liang, “Reviewers’ Identity Cues in Online Product Reviews and Consumers’ Purchase Intention,” *Front Psychol*, vol. 12, Jan. 2022, doi: 10.3389/fpsyg.2021.784173.
 - [53] R. Behbehani, A. Beltagui, M. Schnieder, A. I. Canhoto, and N. Kumar, “Decoupling consumption from value creation: A socio-technical analysis of Everything as a Service (XaaS) adoption by consumers,” in *EurOMA Conference*, 2024.