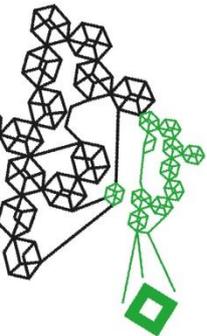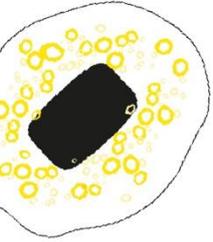# The Scent of Reviews

An Experiment on the effects of Valence and Storytelling on Perfume experience

Manouk Noëlle te Wierik

Faculty of Behavioural, Management, and Social Sciences (BMS)

Communication Science

Dr. R.S. Jacobs

July 21, 2023

Abstract

This paper addresses the field of online consumer reviews in terms of experience goods. Online consumer reviews have been studied numerous times, yet these studies translate to an effect of reviews on beliefs, opinions and attitudes prior to consumption. However, the effect of reviews on the experience and evaluation after consumption remain almost untouched. This research aims to fill this gap. This study aimed to investigate the effects of valence and storytelling, including interaction, on fragrance experience. Furthermore, involvement was incorporated in the analysis as possible moderator. It was expected that positive reviews would have a positive effect of fragrance experience whereas negative reviews would have a negative effect. Besides, storytelling was expected to amplify the effects of valence. Finally, a moderating effect of involvement was expected. A 2x2 experimental design was implemented consisting of storytelling vs non-storytelling along with a positive vs negative review. In this research, a control group was added, resulting in a total of five condition groups and sample size of 137. The finding of this research were argued to not provide enough evidence to support the hypotheses. Thus, all hypotheses were rejected. It is discussed that results might differ for other type of experience goods than perfume. Finally, limitations and recommendations for future research are discussed.

Keywords: Online Consumer Reviews, Fragrance Experience, Storytelling, Valence, Involvement

Introduction

The economic value of experiences have been changing in todays' world of experience and the experience economy. Experiences are seen as an offer on itself and thus distinctive from services, goods and commodities. This is for instance due to the 'inherently personal' attributes as explained by Pine and Gillmore (1998). The characteristics of experience goods cannot be looked up or evaluated prior to consumption as in contrast with search goods. Yet, overlap between these goods could occur. Weathers, Sharma and Wood (2007) sketch an image of a spectrum with search and experience goods on the ends and most products being somewhere in between as they own attributes of both ends. Some even investigate the proposition that the Web and its endless opportunities to search could alter experience goods into search goods (Nakayama, Sutcliff & Wan, 2010). Although this proposition is seemingly insignificant, it is for sure that consumers still try to search for insights prior to consumption, indicating the relevance of reviews.

Within experience goods there are different type of goods. Alt Although there is not a clear classification model on what type of experience good we're dealing with, there is this example of Bhatnagar & Ghose (2004) who did a study on different type of product categories of which recreational and apparel were mentioned as types of experience goods. One can understand that there are differences between e.g. a day to the zoo and a perfume. Goods in the category 'apparel' hit consumers on a personal level.

When relating this to reviews, this fits with the work of Khammash and Griffiths, (2011), which describes the possibilities of reviews to fulfil the search for unique personal experiences. When it comes to an experience good like in this case a perfume, it is extremely hard to imagine what it would smell like. Especially when the scent is new to the consumer. This is therefore a great example of how experience goods bring along uncertainty prior to consumption. In an attempt to take away this uncertainty, reviews are being read by many consumers these days. In this paper there is a specific focus on electronic word of mouth (eWOM). Rosario, de Valck and Sotgiu (2020) grasp the meaning of eWOM in their presented definition: "eWOM is consumer-generated, consumption-related communication that employs digital tools and is directed primarily to other consumers". In other words, online reviews for (prospective) consumers by consumers.

Online consumer reviews have been studied numerous times, yet these studies translate to an effect of reviews on beliefs, opinions and attitudes prior to consumption. However, the effect of reviews on the experience and evaluation after consumption remain almost untouched. This research aims to fill this gap. This study aimed to investigate the effects of valence and storytelling, including interaction, on fragrance experience.

RQ: "Which elements of reviews play an influencing role on fragrance experience?"

Theoretical framework

Valence

The other independent variable, valence, refers to the positive-negative assessment of products or goods. In the domain of online consumer reviews (eWOM), valence has frequently been incorporated in numerous studies. For example, Reinstein and Snyder (2005) demonstrate how a positive review has a positive effect on demand. Likewise, valence has been found to have an influence on trust, brand attitude and purchase intention (Lin & Xu, 2017). This effect might look self-explanatory in many cases, nevertheless, Zou, Yu an Hao (2011) point out that discrepancies in different studies occur and they suggest the moderating role of consumer expertise as explanation.

Furthermore, valence and emotions might occur simultaneously in reviews and could work together. To illustrate: "Positive (negative) affective cues lead to more positive (negative) evaluations and judgments" (Ludwig, 2013). Due to the important role of emotions in the expected effect of storytelling, this could also be applicable to a combination with valence and storytelling as well. In fact, in their study concerning storytelling in reviews, Black and Kelly (2008) looked at a combination of both positive and negative cues in one review. They found that reviews that solely included the positive side of the story were perceived as more helpful as opposed to reviews with a combination. This suggests that positive cues will lead to positive assessments and in turn, negative cues will lead to negative assessments.

Moreover, in relation to odor perception, there seems to be a cognitive bias (positive/negative bias). As Greenberg et al. (2013) state: "Assigning negative bias to an odor prior to an exposure results in the reporting of significantly more health-related symptoms following exposure." Although this study is mainly focussed on the medical world instead of cosmetically perfumes, it has definitely to do with the effects of expectations and beliefs etc. on how an odor will be experienced. It even goes that far that when participants believe they are about to smell something harmful for them, they in fact do experience higher levels of illness. (Greenberg et all, 2013)

Thus, taking into account all the aforementioned, an effect of valence is expected on fragrance experience. The accompanying hypothesis are formulated as followed:

H1a: Positive reviews will have a positive effect on fragrance experience.

H1b: Negative reviews will have a negative effect on fragrance experience.

Storytelling

With its rich history and ancient existence, storytelling remains important in our communication and has evolved to occur in all shapes and sizes. Whereas some scholars stick with the traditional view on storytelling in a purely oral form, others embrace contemporary developments in how stories can be passed along and hence include e.g. textual forms as well (Anderson, 2010). In addition to this disagreement, the many different forms of storytelling makes an all-encompassing definition complex.

This great variety and complexity ensures a broad field of research and as Van de Mieroop (2020) demonstrates, many different elements or features have been analysed and categorised in so-called narrative genres. Some relevant genres in terms of eWOM are for instance small stories, digital narratives or online genres (Georgakopoulou, 2017; De Fina, 2017; Vasquez, 2012). Yet, even when we look at a specific genre, it remains rather difficult to subtract a clear definition to apply, due to the fact that for instance 'researchers often find that people's everyday storytelling consists of a mix of various genres' (Van de Mieroop, 2020).

Furthermore, every form of storytelling arguably has different affordances and restraints. Basically, online consumer reviews (eWOM) allow for a certain setting in which consumers tell their story about their *personal experience*. However, as Hamby, Daniloski and Brinberg (2015) point out, consumer reviews are bound to a smaller magnitude in terms of content and context. Besides distinctions in possibilities and limitations of forms, scholars have paid attention to matters like accomplishments of stories in certain settings (Van de Mieroop (2020).

Proceeding with the latter, one example of a more straightforward accomplishment of storytelling in online consumer reviews is that it could fulfil readers in their search for an unique customer experience from other non-experts (Khammash & Griffiths, 2011). It is interesting that consumers are looking for personal experiences of others in reviews. This notion is also supported in a research conducted by Black and Kelly (2008) which demonstrates that the presence of good story components in a review will lead to higher perceived helpfulness of the reviews. In turn, it is known that higher perceived helpfulness of online reviews can affect purchase behaviour (Ludwig, 2013).

With such findings, it is pertinent to inspect the described components of a good story in a review, especially seen the fact that an all-encompassing definition is complex as described above. Elements described as good story components are trust, character development, tone of superiority, detail and inspiration. (Black & Kelly, 2008). For further elaboration and implementation of these components in this research, see 'Manipulation of storytelling' in the method section.

Storytelling is not only informational, it can convey emotions, strong associations, thoughts and experiences in a memorable way. Or how Karampournioti and Wiedmann (2022) mentioned: "Stories make ideas stick, and they help persuade". It is known that storytelling has persuasive characteristics and it is therefore not surprising that this effect occurs within reviews as well. In a research on the persuasion of narratives in reviews, Hamby et al. (2015) found that reviews that tell a story result in higher purchase intentions.

However, even more interesting is the role of emotions. As stories within reviews are highly personal experiences, emotions come very often, if not always, into play. An interesting thing about emotions is that they can be contagious as explained by the theory of emotional contagion. Herrando et al.(2022), demonstrate that this contagion takes place in online consumer reviews as well. Similarly, Ludwig et al. (2013) state that emotional cues evoke emotional responses, which happens immediate and effortless. Additionally, this affected emotional state leads to change in e.g. attitudes and intentions (Murphy et al., 2013).

Besides the influence of emotional state on matters like beliefs attitude and behaviour, it might be expected that it could even have an influence on the consumer experience of in this case odor perception. As Sucker et al. (2008), mention that in some cases, odor perception is 'guided by use of information in memory, expectations or emotional states'. Meaning that an emotional state prior to the exposure of an olfactory cue, has an influence on the odor

perception. Similarly, in another study participants who were in a negative emotional state (anxious) reported aversion towards a harmless odor (Krusemark et all, 2013).

Although some parts of this study remain exploratory, when taking into account the aforementioned, an influencing effect of storytelling on fragrance experience is expected. Due to the nature of storytelling, a review with storytelling elements is reasonably not plausible to occur. Moreover, as storytelling can strengthen emotions, a review whereby storytelling is included which contains a positive story with positive emotions and associations is expected to have a positive effect on the fragrance experience as opposed to the negative version of such a storytelling induced review. Therefore, in this research, the effect of storytelling is expected to be dependent on valence. Thus, the following hypothesis has been formulated as followed:

H2: The presence of storytelling in a review will amplify the effect of valence as compared to the absence of storytelling in a review.

The role of involvement

The effect of reviews on attitude and purchase intention has been extensively studied. When looking at attitude and how this can be influenced, the Elaboration Likelihood Model (ELM) could serve as explanation. This model is based on the notion that persuasion plays a key role in influencing attitude. It distinguishes how persuasion happens based on the level of involvement of the receiver (UT, 2003/2004). In a study concerning eWOM and consumer intention, Park and Lee (2008) state that consumer involvement is a key moderator on the processing of the review. They explain that consumers with higher levels of involvement are more likely to take in more extensive content of the persuasion, whereas consumers with lower levels of involvement are more prone to superficial cues of the persuasion.

In a 'normal' setting those reading reviews are typically deliberately looking for it themselves, indicating a high level of involvement (motivation and willingness to digest the content). However, due to the nature of this experimental research some participants might have a low level of involvement. Therefore, the differences in level of involvement could be incorporated as possible moderator.

Relating this to the current context, as mentioned before, storytelling has proven to have persuasive characteristics. Additionally, in this research, the storytelling induced manipulations have more extensive content as will be elaborated in the method section. In other words, consumers with high involvement would be more prone to the effects of storytelling compared to consumers with low involvement. On the contrary, in this research, valence has more superficial cues of the persuasion, which would translate to the suggestion that consumers with low involvement would be more prone or sensitive to the effects of valence compared to consumers with high involvement. Therefore, the following hypotheses have been formulated;

H3a: A high level of involvement strengthens the effect of storytelling as compared to a low level of involvement.

H3b: A low level of involvement strengthens the effect of valence as compared to a high level of involvement.

Method

As the main focus in this experimental research is to explore the effects of two independent factors, i.e. review valence and storytelling, on fragrance experience between subjects, a quantitative 2x2 factorial design was applied. Hereby, the factor 'review valence' is divided in positive and negative, whereas the factor 'storytelling' is divided in on the hand the presence of storytelling in a review, referred to as storytelling, and on the other hand the absence of storytelling in a review, referred to as non-storytelling. In addition, a control group was included, resulting in a total of five condition groups (Table 1).

Moreover, this design allows to test for an interaction effect between these two factors as well as investigate the effect of involvement in order to assess the aforementioned hypotheses. Regarding the data analysis methods, the statistical data analysis tool focussed on social sciences called 'SPSS' was used.

Table 1

Representation of the experimental conditions

Condition group	Review Valence	Presence of Storytelling
1) Positive Storytelling	Positive	Storytelling
2) Negative Storytelling	Negative	Storytelling
3) Positive Non-Storytelling	Positive	Non-Storytelling
4) Negative Non-Storytelling	Negative	Non-Storytelling
5) Control group	-No review	-No review

Procedure

The data was collected through an online questionnaire constructed with the survey software Qualtrics. Although the respondents and the researcher were physically present, an online questionnaire was chosen to represent the online shopping environment during the exposure of the stimuli. Participants received explanation on the research in person and even though there was an oral agreement to participate, the first page of this questionnaire served as an informed consent form nevertheless.

Proceeding with the questionnaire, the first questions concerned demographics, followed with some statements about their own involvement with perfume(s). Secondly, the respondents were asked to read the stimulus presented to them based on their condition group. This allocation was executed by Qualtrics through randomization. It was measured how long the respondents took to read, that is to say how long it took to move to the next page. On this new page, questions were asked about their feelings and interest in this perfume after reading the review.

After that it was time to test the actual perfume. This was followed by multiple questions and statements about their fragrance experience; how they feel, whether their expectations were met, their evaluation of the perfume, their purchase intention and their rating. Elaboration of the statements and their corresponding constructs can be found under 'Measurements'.

Stimulus Material

In order to measure the influence of valence and storytelling in a review, four reviews have been designed with these manipulations incorporated. Besides the changes in manipulations, the reviews and their contents are made similar. For instance, in order to get closer to reality, the looks of the reviews have been improved by adding a photo and a name of the reviewer. These visualisations are the same in all the reviews.

Besides the review, the environment on which the stimuli was shown was altered to reality as well. In a normal setting it would not occur often that one reads a review without having seen any information of the product. Therefore, a visualisation of an online webpage was created which showed little information of the product (Appendix A, Figure A.1).

Apart from this, the review was shown on a screen as the questionnaire was online, thereby getting closer to the online environment. Moreover, hereby it was accounted for the source of the review, as it might influence whether the review comes from the website itself or from an external third-party site. Thus, the information from the website and the review were showed separately.

It should also be noted that a review is a source of information. Therefore if the control group would not get to see anything, there would not only be the differences in the absence of the targeted manipulations, but also the absence of information. In this case, the informational

value is designed by focussing on ingredients of the perfume, as the reviews also refer to this. To make sure this variable is fixed for all groups, all condition groups got to see the informational page, including the control group.

Manipulation of storytelling

The absence or presence of storytelling in a review was manipulated based on good story components as described in the theoretical framework, which consisted of multiple components. 1) 'Trust' can be established by sharing personal information about for instance their family; 2) 'Character development or establishment' is best done at the beginning of the review due to the short amount of 'time' in a review; 3) 'Tone of superiority' should be incorporated by the absence of it, as the reviewer should be superior to the reader; 4) 'Detail' is done with sensory details and emotional content; 5) Besides, a good story evokes feelings (Black & Kelly, 2008).

Besides these differences, other factors were displayed similar. For example, in a review with storytelling it is being told how the reviewer tried something new to surprise her husband, whereas in the review without storytelling this is altered to how the reviewer tried something new. Due to the nature of storytelling it was inevitable that these reviews were a bit longer, yet still similar. The manipulations for storytelling are shown in Appendix A (Figure A.2 to Figure A.5).

Manipulation valence

The manipulation of valence was relatively simple. In general, a positive word within a positive review is replaced with a negative word for the negative review. Thereby it was taken into account that there is a similarity in e.g. tone and meaning or intensity. The sentence structures remained unchanged to ensure the same context in the different groups. For instance, the sentence 'I am impressed' was changed to 'I am not impressed'.

Finally, in order for the reviews to appear more realistic, a visualization of the evaluation was included. This was done by means of a green thumbs up for the positive reviews and a red thumbs down for the negative reviews, thereby eliminating the elements of numbers in e.g. stars or ratings. The manipulations for valence are shown in Appendix A (Figure A.2 to Figure A.5).

Pre-test

The stimulus material as well as the questionnaire were tested by means of a small qualitative pre-test. The aim of this pre-test was to discover issues or difficulties in language, tone, structure, comprehension and to receive feedback. Participants were selected based on the personal network. As the sample of the main study will allow for a wide variety within age, it has been decided to select different ages. A total number of four participants participated in the pre-test. The age of the participants ranged from a minimum of 22 years old to a maximum of 82 years old.

The procedure of the pre-test started with an explanation of the aim of this test. Subsequently, participants were asked to read all of the reviews by means of a think-aloud-protocol. In addition, some questions were posed regarding specific words or sentences, as well as follow up questions based on their answers. It was for instances asked with which word they would replace a certain difficult word. Finally, contents of the questionnaire itself was also discussed.

The results showed that some words in the reviews were too complicated. As an example, this was the case for the word 'affection'. Therefore, adaptions were made accordingly. Besides that, this pre-test resulted in simplified sentence structure. Additionally, when it comes to the questionnaire, some words within the pairs of adjectives of the construct 'fragrance evaluation' have been adjusted. Not all given considerations were taken into account. It was for instances noted that some ingredients were not understood, however as this is due to knowledge rather than a flaw in language for instance, it was decided to leave this unchanged.

Respondents

As for the sampling method, the non-probability 'convenience sampling' has been applied. To gather respondents, passengers on the streets in the city centre of Enschede in the Netherlands were asked to participate in the study. This location was appropriate in terms of crowdedness while at the same time being spacious enough to feel calm which minimized distracting factors. Additionally, an advantage of this location is the open air which provides for risk reduction in terms of sensitivity for the perfume.

Due to the odds of a relatively large number of Dutch speakers in this location, it was decided to alter the language of the questionnaire to Dutch. Consequently, one of the inclusion criteria was being able to read and understand Dutch. Furthermore, there was an age restriction of a minimum of 16 years old. Finally, to leave out gender as another possible influencing factor, it has been decided to only include females.

Sample characteristics

In total, the questionnaire had been started 146 times, however, not all of these resulted in completely finished responses. It has been decided to remove the incomplete responses (n=5) based on a set criteria of a progress completion of at least 80%. Furthermore, it occurred that a couple of males filled in the questionnaire as well (n=4). Due to the exclusion criteria as mentioned above, these responses were removed from the dataset. Thus, this leaves us with a sample size of N=137.

The age of the participants ranged from a minimum of 16 years old to a maximum of 79 years old, with an average age of 31 years old (SD=17.23). The total time participants have spent on the entire questionnaire was on average 532,59 seconds (SD=308.73) which translates to almost 9 minutes and the amount of time they have spent on the review was on average 53,24 seconds (SD=50.00). Finally, participants in this sample were characterised by their level of involvement in perfumes with an average of 5.29 (SD=1.37), This was measured on a 7-point scale whereby a higher number represents a higher level of involvement.

Distribution of sample characteristics

As mentioned before, the respondents were allocated to one of the condition groups by means of randomization. In order to assess the randomization and check whether this ensured an even distribution, the variances have been analysed for the variables age, the time it took to complete the questionnaire, the time spent to look at the presented stimuli and the level of involvement. A one way ANOVA has been used to compare the means among the different condition groups of each of the abovementioned variable. Table 2 underneath gives an overview over the means of each variable per condition group.

The analysis shows that there are no significant differences between the five condition groups for age $F(4,124)=0.246$, $p=.911$ ($\geq .05$). This was also the case for 'duration of questionnaire' $F(4,132)=0.629$, $p=.643$ ($\geq .05$) as well as for level of involvement $F(4,132)=0.974$, $p=.424$ ($\geq .05$). this implies an equal distribution.

However, regarding the time that is spent to look at the presented stimuli, the analysis shows a significant result for differences between the condition groups $F(4,132)=2.89$, $p=.025$ ($\leq .05$). This is interesting as this is not a typical variable used to characterise a sample by. On the one hand, in this case this analysis is used to inspect whether there is a group in which participants passed through the stimuli remarkably fast. On the other hand, this analysis serves as a test for the manipulations as well, whereby logically seen, a longer review will take longer to read. However, as explained before, the length of the review is not an intended factor so preferably there would be no significant differences here.

To further inspect this significance and to prevent a Type I Error, a post-hoc test has been included in this analysis by means of Bonferroni. When comparing every individual condition group with one another, it can be seen that there is only a significant difference between the control group who got to see only the limited information of the perfume and the group who got to see the negative review with storytelling elements ($p=.015$). This analysis also implies that the differences between all the other groups that are found in this study are relatively small and are therefore disregarded. Thus, there are no significant differences between the four manipulations of the reviews.

Table 2

Distribution of sample characterics

Condition group	Participants	Age		Duration of questionnaire		Time spent on stimuli		Level of Involvement	
Condition	N	M	SD	M	SD	M	SD	M	SD
1) Positive Storytelling	26	30.58	15.66	561.27	326.95	49.72	21.84	4.98	1.36
2) Negative Storytelling	27	28.15	16.99	570.56	306.77	73.98	80.39	5.53	1.32
3) Positive Non-Storytelling	29	32.04	18.17	568.62	337.25	61.19	59.47	5.05	1.50
4) Negative Non-Storytelling	28	31.26	17.64	483.79	290.14	49.59	28.79	5.42	1.19
5) Control group	27	32.50	18.44	478.93	287.13	31.13	20.85	5.50	1.47
Total	137	30.91	17.23	532.59	308.73	53.24	50.00	5.29	1.37

Measurements

To adequately measure the effect of review valence and storytelling on the experience of perfume, multiple questions were formulated, whereby a total of five constructs were build. A prominent factor in measuring the hypotheses is perfume experience, which can consists of multiple elements. To fit the comprehensiveness of this factor, the construct 'fragrance evaluation' is more elaborated. However, questions like 'How would you rate this perfume from zero to then?' are of course also a form of evaluation. Likewise, purchase intention can also be seen as an evaluation.

Furthermore, 'interest' has been added for the questionnaire to represent more realistic and complete as well as to allow for a variable to be tested on the moment after reading the review yet before testing the perfume. As described in the procedure, some other questions were asked as well, as for instance demographic questions including age and gender. Thus, these variables are all included to assess whether and to what extent these various components are influenced by the independent variables to better answer the research question.

Furthermore, all items within the constructs were measured using a 7-point Likert scale, ranging from 'completely disagree' to 'completely agree'. Finally, all constructs were tested for validity and reliability as explained below.

Involvement

As described in the theoretical framework, the measurement of involvement is essential to test hypothesis H3a and H3b. This construct consists of six items. These are based on the work of Zaichkowsky (1985), in which the construct successfully meets various standards of validity and reliability. Nevertheless, the items were slightly adjusted to this case. For instance, the item 'I would be interested in reading the *Consumer Reports* article about this product' was changed to 'I am interested in reviews of others about perfume'. In addition, to account for the high favourability within perfumes, the item 'I am open to discover new perfumes' was added.

Interest

The creation of items in this construct are two-folded. On the one hand there are self-evident items as the rather obvious item 'I am interested in this perfume'. On the other hand, items were inspired by Horn and Salvendy (2007) who discussed various aspects which had to do with e.g. behavioural intention and attitude. An item that evolved from this is for instance 'It is likely that I will actively search for this perfume'. Furthermore, another part of interest has to do with aesthetics and desire of the product (Horn & Salvendy, 2007). This resulted in for instance the item 'This perfume appeals to me'. In total, this construct consists of six items.

Expectations-met scale

The items in this construct were based on the construct used by Meng, Zamudio and Jewell (2021) to measure the differences between expected and experienced scents. The items have been adapted as the items in their paper were measured by means of pairs of adjectives, whereas in this study it is measured by means of statements. An example of a statement is the following; 'The actual fragrance is similar to what I imagined'. This construct consists of 4 items, of which one item was recoded after data collection.

Fragrance evaluation

The experience of a perfume could have various aspects. To fit the comprehensiveness of this factor, multiple components have been included, yet not all of them. For example, when thinking about how a perfume smells, one might think about the ingredients or aspects like whether it is sweet, floral or woody. However, this does not say anything about how the perfume is perceived in terms of (un)favourability. For some a sweet perfume would be seen as a benefit or enjoyment, while for others this is the opposite. Therefore, it has been decided to leave this type of judgements out to fit the research objectives better.

The items in this construct were based on multiple papers on olfactory experiences. These papers elaborated on various aspects of these experiences as for instance intensity, pleasantness, sensuality, hedonic performance, sensorial aspects et cetera (Barwich, 2014; Spears & Singh, 2004; Muniesa & Trébuchet-Breitwiller, 2010; Porcherot et al., 2010). All in all, this resulted in a relatively big construct with 16 items. Therefore, the items were

measured by means of pairs of adjectives because this is known to be easy, quick and effective.

Purchase intention

Lastly, this fifth construct is a construct that has been widely used in previous studies. Besides that, the creation of these items seem obvious as for example the item 'I would buy this perfume'. Nevertheless, this construct was based on previous studies like those of Spears and Singh (2004) as well as Meng, Zamudio and Jewell (2021). This construct consists of six items in the form of statements.

Validity test

A confirmatory Factor Analysis has been applied to test for the validity of the constructs. That is to say, it has been controlled whether the items that were combined in a group measure the same thing. The Kaiser-Meyer-Olkin Measure was used to control for the suitability of the data for a Factor Analysis (KMO=0.903 $\geq$ 0.50). Additionally, by means of the principal components method, the fixed number of factors to extract was set at 5, where after it was shown that the 5 factors explain 69,7% of the total variance. Furthermore, small coefficients were suppressed below 0.40.

When looking at the rotated component matrix (varimax), three matters were remarkable. Firstly, one item within the construct 'fragrance evaluation' did not score high enough and did therefore not appear in any of the five factor columns. This suggestion to leave out the item was followed and thus, the item 'weak/strong' was removed from the construct. Secondly, another item within 'fragrance evaluation' appeared to be matched with other items of the construct 'purchase intention' as well. Though, as it has a higher score in the column of the other items of 'fragrance evaluation', it has been decided to leave this unchanged.

Thirdly, all items within 'purchase intention' were also shown in the same factor column including items of 'fragrance evaluation'. Similarly to above, these two constructs are matched. This means that these two constructs measure similar things, which is not crazy as they are both evaluative. However, there was still a clear difference in the magnitude of the values, namely all below 0.5 when matched with 'fragrance evaluation' compared to all above

0.7 when matched with 'purchase intention'. Therefore, based on the factor analysis, no adjustment where made to this construct.

The rest of the constructs were indeed behaving as intended and no adjustments were suggested to be made. Subsequently, the item reliability was tested.

Reliability test

To further asses the constructs, an item reliability analysis was used by means of the Cronbach's alpha, whereby a high alpha represents a high reliability. The analysis showed high values of Cronbach's Alpha, representing good and excellent scores of internal consistency.

However, for the construct 'expectations met' it was shown that removing one item would lead to a relatively improvement for the reliability. This was executed and after that, the construct consisted of 3 items. For the other constructs, no adjustments were made. All Cronbach's Alpha of the final states of the constructs are displayed in Table 3 below.

Table 3

Cronbach's Alpha per construct

Name of construct	Items (N)	Cronbach's Alpha (α)
Involvement	6	.853
Interest	6	.930
Expectations met	3	.878
Fragrance evaluation	15	.957
Purchase interest	6	.975

Results

As discussed before, to better asses the research question and the formulated hypotheses, multiple measurement constructs have been included as dependent variables. Therefore, an analysis for general linear model with multiple variables was chosen. First, the main effects will be analysed after which the proposed moderating effect of involvement will be analysed as well. In both cases a multivariate analysis of variance, hereafter referred to as 'MANOVA', was conducted.

In Table 4 below, the mean scores on the dependent variables can be compared for the reviews in which storytelling was incorporated as opposed to the reviews that did not incorporate this as well as opposed to the control group. Similarly, the mean scores on the dependent variables can be compared for the reviews that were positive as opposed to those that were negative as well as opposed to again the control group.

Table 4

Descriptive statistics of the dependent variables for each manipulation

Dependent variable	Absence or presence of storytelling		Review valence		Control group
	Storytelling	Non-Storytelling	Positive	Negative	Information only
Interest	M=3.97 SD=1.51	M=4.11 SD=1.32	M=4.15 SD=1.44	M=3.94 SD=1.38	M=4.44 SD=1.17
Expectations met	M=3.81 SD=1.38	M=3.77 SD=1.58	M=3.78 SD=1.31	M=3.81 SD=1.64	M=4.23 SD=1.56
Fragrance evaluation	M=4.08 SD=1.50	M=4.21 SD=1.31	M=4.07 SD=1.28	M=4.22 SD=1.52	M=4.66 SD=1.36
Purchase intention	M=2.73 SD=1.67	M=2.87 SD=1.61	M=2.55 SD=1.57	M=3.05 SD=1.67	M=3.30 SD=1.75

Main effects

In order to test the main effects of valence and storytelling, including the interaction effect, a 'MANOVA' as mentioned before was conducted. More specifically, the included dependent variables in this analysis are 'interest', 'expectations met', 'fragrance evaluation' and 'purchase intention'.

This analysis showed that both storytelling ($F(4,128)= 0.116$, $p=.977 \geq .05$; Wilk's Λ =0.996, partial $\eta2 = .004$) and valence ($F(4,128)= 1.794$, $p=.134 \geq .05$; Wilk's Λ =0.947, partial $\eta2 = .053$) were insignificant in this model. In addition, when we look at the interaction effect between storytelling and valence, the analysis showed again an insignificant effect ($F(4,128)= 0.647$, $p=.630 \geq .05$; Wilk's Λ =0.980, partial $\eta2 = .020$).

Besides these general effects, the tests of between-subjects effects gives more insight on the effects on each dependent variable separately. In line with the multivariate test, the tests of between-subjects effects showed no significant results as can be seen in Table 5 below.

Table 5

Tests of Between-Subject Effects

		F-value	*Sig.*
Storytelling			
	Interest	0.241	.624
	Expectations met	0.017	.897
	Fragrance evaluation	0.244	.622
	Purchase intention	0.215	.644
Valence			
	Interest	0.546	.461
	Expectations met	0.022	.882
	Fragrance evaluation	0.327	.569
	Purchase intention	2.510	.116
Storytelling * Valence			
	Interest	0.284	.595
	Expectations met	1.415	.236
	Fragrance evaluation	0.054	.816
	Purchase intention	0.085	.771

Interaction effect involvement

In order to assess the proposed moderating effect of involvement, another MANOVA was conducted. More specifically, the independent variables storytelling and valence were included as well, yet in this analysis, involvement was added as covariate. Moreover, again the included dependent variables in this analysis are 'interest', 'expectations met', 'fragrance evaluation' and 'purchase intention'.

The variable involvement as measured by means of the construct was transformed for the purpose of the analysis and to better fit the hypothesis. This resulted in a nominal variable with the categories low involvement and high involvement. This was executed based on the median and it was accounted for the frequency distribution as derived from a histogram.

This time, the analysis showed statistical significance for valence in this model ($F(4,123)$= 2.634, p=.037 ≤ .05 ; Wilk's Λ =0.921, partial η2 = .079), yet not for storytelling ($F(4,123)$= 0.685, p=.603 ≥ .05 ; Wilk's Λ =0.978, partial η2 = .022) or the interaction effect of storytelling and valence ($F(4,123)$= 0.132, p=.970 ≥ .05 ; Wilk's Λ =0.996, partial η2 = .004). Furthermore, involvement showed no significant effect on either storytelling ($F(4,123)$= 1.379, p=.245 ≥ .05 ; Wilk's Λ =0.957, partial η2 = .043) or valence ($F(4,123)$= 0.716, p=.583 ≥ .05 ; Wilk's Λ =0.977, partial η2 = .023), neither on the interaction between storytelling and valence ($F(4,123)$= 0.228, p=.922 ≥ .05 ; Wilk's Λ =0.993, partial η2 = .007).

Besides looking at the significances of p-values, the Partial Eta Squared can also be compared. A Partial Eta Squared of 0.01 indicates a small effect whereas 0.06 indicates a medium effect and 0.14 a large effect. Valence stood out with a Partial Eta Squared of 0.079, indicating a medium effect.

Besides these outcomes of the multivariate tests, the tests of between subjects effects gives more insight on the effects on each dependent variable separately. This shows that there are no significant results for none of the independent variables on any of the dependent variables. This means that also for valence no significant result is shown, despite the significant result of before. The outcomes on the tests of between-subjects effect of this analysis can be seen in Table 6 below.

Table 6

Tests of Between-Subjects Effects

		F-value	*Sig.*
Storytelling			
	Interest	1.856	.176
	Expectations met	0.134	.715
	Fragrance evaluation	0.049	.825
	Purchase intention	1.303	.256
Valence			
	Interest	2.149	.145
	Expectations met	0.796	.374
	Fragrance evaluation	0.208	.650
	Purchase intention	1.638	.203
Storytelling * Valence			
	Interest	0.238	.626
	Expectations met	0.086	.770
	Fragrance evaluation	0.003	.954
	Purchase intention	0.108	.743
Valence * Involvement			
	Interest	0.005	.945
	Expectations met	0.492	.485
	Fragrance evaluation	0.125	.725
	Purchase intention	1.079	.301
Storytelling * Involvement			
	Interest	2.197	.141
	Expectations met	0.222	.638
	Fragrance evaluation	0.000	.988
	Purchase intention	2.690	.103
Storytelling * Valence * Involvement			
	Interest	0.001	.978
	Expectations met	0.027	.869
	Fragrance evaluation	0.147	.702
	Purchase intention	0.474	.492

Discussion

This study aimed to investigate the effects of valence and storytelling, including interaction, on fragrance experience. Furthermore, involvement was incorporated in the analysis as possible moderator. It was expected that positive reviews would have a positive effect of fragrance experience whereas negative reviews would have a negative effect. Besides, storytelling was expected to amplify the effects of valence. Finally, a moderating effect of involvement was expected.

After conducting two multivariate analysis of variance, it was shown in the results that there were barely any significances. Other than the significant effect of valence in the linear model whereby involvement was included, there were no other significances. Furthermore, it is worth mentioning that this significant effect is a statistical significance. It helps to evaluate whether a null-hypothesis can be rejected, meaning that the chances of seeing differences within the population are high. However, these results still concern chances.

Also, the one reported significant p-value is on the higher side of the general consensus criteria of 5%. Besides, as this significance was later on in the analysis not seen in the tests of between-subjects effects, it is not respected as a strong effect. Therefore, it is in this case suggested that the H01 will not be rejected, and therefore H1 is not supported. Thus, none of the hypotheses were supported in this research.

Implications

As the results were not in line with the expected outcomes, this could have various implications. Firstly, it could be that these results objectify the theory behind the expected outcomes, that is to say, that this research would serve as evidence to contradict those theories. However, as this study is not a one on one replication of a research or a theory, but rather a combination of various studies and theories, this is deemed unlikely.

Secondly, it could be that the expected outcomes were based on research of a different context than the context of this performed study. This is presumably a justifiable explanation. As this study aimed to fill a research gap, it is rather complex to decide when a context is similar or fitting enough. However, these results could also imply that the effects of reviews are not as prominent for perfume as studied in this research compared to other type of

products as studied in previous research. Seeing the results, a practical implications here could be that reviews appear to not have an effect on how the perfume is experienced, but neither on interest or purchase intention.

Limitations and further research

Undeniably, it is important to discuss limitations of the research. Firstly, in order to for the results of this research to be applicable on the population, it should be a good representation of reality. This is where this research has a shortcoming. For instance, as some point, despite the fact the research was conducted in the open air, the scent of the perfume became more persistent to a degree where it was possible to detect the perfume already, even before reading the review. Furthermore, it had not been included in the questionnaire whether the participants were already familiar with the specific perfume. Logically, this could be an influencing factor for the results.

On another note, a limitation in this research is the absence of a manipulation check in the sense of a straightforward question to assess whether the participants perceived the elements of the manipulation as intended. There might be a possibility that participants did not perceive any difference between the manipulations for instance or that other factors besides those intended were perceived as well. As example for the latter, it could be that e.g. the length of a review plays a role. Finally, it should be noted that this study was performed with a relatively small sample size, which limits the findings. Logically following on these limitations, it is recommended for future research to take this into consideration.

Conclusion

This experimental study had a main focus on the effects of review valence and storytelling, including an interaction effect, on fragrance experience in general. However, involvement was incorporated as well as possible moderator. Furthermore, multiple dependent variables were investigated, which were the following; 'interest', 'expectations met', 'fragrance evaluation' and 'purchase intention'. A 2x2 experimental design was implemented consisting of storytelling vs non-storytelling along with a positive vs negative review. In this research, a control group was added, resulting in a total of five condition groups. The main effects were analysed, after which the proposed moderating effect of involvement was analysed as well. In

both cases, this was done by means of a multivariate analysis of variance. Based on the findings of this research, the hypotheses were not supported. This suggests there is no effect for review valence and storytelling, including an interaction effect, as well as that there is no effect for involvement. Nevertheless, this study is an addition to insights on the effects of reviews in the specific case of perfumes.

Reference list

Anderson, K.E. (2010). Storytelling

Barwich, AS. (2014). A Sense So Rare: Measuring Olfactory Experiences and Making a Case for a Process Perspective on Sensory Perception. Biol Theory 9, 258–268. Doi:.org/10.1007/s13752-014-0165-z

Black & Kelly (2008). A storytelling perspective on online customer reviews reporting service failure and recovery

Greenberg et all (2013). The perception of odor is not a surrogate marker for chemical exposure: a review of factors influencing human odor perception

Hamby et all (2015). How consumer reviews persuade through narratives

Herrando et all (2022). Emotional contagion triggered by online consumer reviews: Evidence from a neuroscience study

Horn, D. & Salvendy, G. (2006). Product creativity: conceptual model, measurement and characteristics, Theoretical Issues in Ergonomics Science, 7:4, 395-412, DOI: 10.1080/14639220500078195

Khammash & Griffiths, (2011). 'Arrivederci CIAO.com, Buongiorno Bing.com'—Electronic word-of-mouth (eWOM), antecedences and consequences

Karampournioti & Wiedmann (2022). Storytelling in online shops: the impacts on explicit and implicit user experience, brand perceptions and behavioral intention

Krusemark et all (2013). When the sense of smell meets emotion: anxiety-state-dependent olfactory processing and neural circuitry adaptation

Lin & Xu (2017). Effectiveness of online consumer reviews: The influence of valence, reviewer ethnicity, social distance and source trustworthiness

Ludwig et al. (2013). More than Words: The Influence of Affective Content and Linguistic Style Matches in Online Reviews on Conversion Rates

Meng, Zamudio & Jewell (2021). What's in a name? Scent brand names, olfactory imagery, and purchase intention.

Muniesa, F. & Trébuchet-Breitwiller, AS. (2010). Becoming a measuring instrument, Journal of Cultural Economy, 3:3, 321-337, DOI: 10.1080/17530350.2010.506318

Murphy et al. (2013). Narrative Versus Nonnarrative: The Role of Identification, Transportation, and Emotion in Reducing Health Disparities

Park & Lee (2008). eWOM overload and its effect on consumer behavioral intention depending on consumer involvement

Pine & Gillmore (1998). Welcome to the Experience Economy

Porcherot, C., Delplanque, S., Raviot-Derrien, S., Le Calvé, B., Chrea, C., Gaudreau, N. & Cayeux, I., (2010). How do you feel when you smell this? Optimization of a verbal measurement of odor-elicited emotions

Reinstein & Snyder (2005). The influence of expert reviews on consumer demand for experience goods: a case study of movie critics

Spears, N., & N. Singh, S.N. (2004) Measuring Attitude toward the Brand and Purchase Intentions, Journal of Current Issues & Research in Advertising, 26:2, 53-66, DOI: 10.1080/10641734.2004.10505164

Sucker et al., (2008). Odor frequency and odor annoyance. Part I: assessment of frequency, intensity and hedonic tone of environmental odors in the field

UT (2003/2004). Communication theories

Van de Mieroop (2021). The Narrative Dimensions Model and an exploration of various narrative genres

Vasquez (2012). Narrativity and involvement in online consumer reviews: The case of TripAdvisor

Zaichkowsky, J.L. (1985). Measuring the Involvement Construct. Journal of Consumer Research Vol. 12, No. 3 pp. 341-352

Zou, Yu & Hao (2011). Does the Valence of Online Consumer Reviews matter for Consumer Decision Making? The Moderating Role of Consumer Expertise

Appendices

Appendix A: Stimuli

Figure A.1

Visualisation informational page, fixed stimuli for all five condition groups.

Home > Merken > Alyssa Ashley

Alyssa Ashley Musk Eau de Toilette

Productinformatie

Eau de toilette. Een zowel bloemig als houtachtig parfum, met in de hoofdrol de aroma's van muskus.

Topnoten: Siciliaanse bergamot
Hartnoten: Iris, tonkabonen, muskus, vanille
Basisnoten: Roos, jasmijn, ylang ylang, geranium

Visualisation of condition group 1) Postive Storytelling

Charlotte Janssen
2 dagen geleden

Ik ben blij verrast!

Blij dat ik dit parfum heb gekocht! Ik had al tijden geen parfum meer gedragen en ik wou graag m'n man ermee verrassen. De romantiek miste de laatste jaren, druk met werk en de kids.. Ik zocht iets om uit die sleur te komen, iets nieuws, iets anders. Nou, dit parfum is zeker anders. Er worden veel verschillende type geuren gecombineerd, die gewoon echt fijn samenwerken. De aroma's van muskus en tonkabonen samen met de zoetere aroma's van vanille en jasmijn, zoiets had ik nog niet eerder geroken. Ik vind het wel sexy, het ruikt nieuw en verfrissend.

Ikzelf was onder de indruk, maar ik was benieuwd wat m'n man vond. Ik droeg het tijdens 't etentje voor ons jubileum. "Wat ruik ik toch?", vroeg hij glunderend. Hij vond het lekker, ik voelde me opgewekt, plan geslaagd! Ik moest denken aan vroeger, toen alles vanzelf ging. Nu was dat eigenlijk net zo. Die aandacht, vreugde en liefde zoals toen was er weer. Die charme is precies wat er miste en dit parfum helpt er zo bij! Wat lekker ruiken wel niet met m'n zelfvertrouwen doet. Ik voel me zelfs aantrekkelijk en zelfverzekerd.

Ik had dit parfum eerder moeten dragen!

Visualisation of condition group 2) Negative Storytelling

Charlotte Janssen
2 dagen geleden

Ik ben teleurgesteld!

Spijt dat ik dit parfum heb gekocht! Ik had al tijden geen parfum meer gedragen en ik wou graag m'n man ermee verrassen. De romantiek miste de laatste jaren, druk met werk en de kids.... Ik zocht iets om uit die sleur te komen, iets nieuws, iets anders. Nou, dit parfum is zeker anders. Er worden veel verschillende type geuren gecombineerd, die gewoon echt niet fijn samenwerken. De aroma's van muskus en tonkabonen samen met de zoetere aroma's van vanille en jasmijn, zoiets had ik nog niet eerder geroken. Ik vind het allesbehalve sexy, het ruikt oud en muffig.

Ikzelf was niet onder de indruk, maar ik was benieuwd wat m'n man ervan vond. Ik droeg het tijdens 't etentje voor ons jubileum. "Wat ruik ik toch?", vroeg hij walgend. Hij vond het niet lekker, ik voelde me somber, plan mislukt! Ik moest denken aan vroeger, toen alles vanzelf ging. Nu was dat niet zo. Die aandacht, vreugde en liefde zoals vroeger was er niet meer. Die charme is precies wat er mist en dit parfum helpt er niet bij! Wat vies ruiken wel niet met m'n zelfvertrouwen doet. Ik voel me zelfs onaantrekkelijk en onzeker.

Ik had dit parfum nooit moeten dragen!

Visualisation of condition group 3) Positive Non-Storytelling

Charlotte Janssen
2 dagen geleden

Ik ben blij verrast!

Ik heb dit parfum vorige week gekocht en ben onder de indruk! Dit is de eerste keer dat ik dit luchtje uitprobeer, ik wou eens een keer iets nieuws, iets anders. Nou, dit parfum is zeker anders, in positieve zin! Het onderscheidt zich door veel verschillende type geuren te combineren. Los van elkaar zijn het ook al fijne geuren, maar ik denk dat het de combinatie of samenstelling is waardoor het zo goed werkt in dit parfum.

De aroma's van muskus en tonkabonen samen met de zoetere aroma's van vanille en jasmijn, zoiets had ik nog niet eerder geroken. Ik vind het wel sexy, het ruikt nieuw en verfrissend, gewoon echt lekker! Precies wat je wilt in een parfum, ik raad het zeker aan. Goed ruiken heeft nog nooit iemand slecht gedaan. Dit parfum geeft me een zelfverzekerd en aantrekkelijk gevoel.

Ik had het eerder moeten kopen!

Figure A.5

Visualisation of condition group 4) Negative Non-Storytelling

Charlotte Janssen
2 dagen geleden

Ik ben teleurgesteld!

Ik heb dit parfum vorige week gekocht en ben niet onder de indruk. Dit is de eerste keer dat ik dit luchtje uitprobeer, ik wou eens een keer iets nieuws, iets anders. Nou, dit parfum is zeker anders, in negatieve zin! Het onderscheidt zich door veel verschillende type geuren te combineren. Los van elkaar zijn het misschien wel fijne geuren, maar ik denk dat het de combinatie of samenstelling is waardoor het niet werkt in dit parfum.

De aroma's van muskus en tonkabonen samen met de zoetere aroma's van vanille en jasmijn, zoiets had ik nog niet eerder geroken. Ik vind het allesbehalve sexy, het ruikt oud en muffig, gewoon echt niet lekker! Precies wat je niet wilt in een parfum, ik raad het zeker af. Slecht ruiken heeft nog nooit iemand goed gedaan. Dit parfum geeft me een onzeker en onaantrekkelijk gevoel.

Ik had het nooit moeten kopen.