Surprise … Surprise…

An Empirical Investigation on How Surprise is Connected to Customer Satisfaction

Joëlle Vanhamme

ERASMUS RESEARCH INSTITUTE OF MANAGEMENT

REPORT SERIES
RESEARCH IN MANAGEMENT

BIBLIOGRAPHIC DATA AND CLASSIFICATIONS		
Abstract	This research investigates the specific influence of the emotion of surprise on customer transaction-specific satisfaction. Four empirical studies–two field studies (a diary study and a cross section survey) and two experiments–were conducted. The results show that surprise positively [negatively] influences satisfaction directly and indirectly (via the amplification of positive [negative] emotions), even when disconfirmation is taken into account in the model. The amplification property of surprise and the How-do-I-feel-about-it? heuristic are believed to explain this influence. Some results also show that surprised customers display higher levels of satisfaction and dissatisfaction than non surprised customers.	
Library of Congress Classification (LCC)	5001-6182	Business
	5410-5417.5	Marketing
	HF 5415.32+	Consumer Satisfaction
Journal of Economic Literature (JEL)	M	Business Administration and Business Economics
	M 31 C 44	Marketing Statistical Decision Theory
	M 39 D 12	Marketing and advertising: Other Consumer Economics: Empirical Analysis
European Business Schools Library Group (EBSLG)	85 A	Business General
	280 G 255 A	Managing the marketing function Decision theory (general)
	280 N	Consumerbehavior
Gemeenschappelijke Onderwerpsontsluiting (GOO)		
Classification GOO	85.00	Bedrijfskunde, Organisatiekunde: algemeen
	85.40 85.03	Marketing Methoden en technieken, operations research
	85.40 83.05	Marketing Economische psychologie
Keywords GOO	Bedrijfskunde / Bedrijfseconomie	
	Marketing / Besliskunde	
	Consumentengedrag, Verrassing	
Free keywords	surprise, satisfaction, disconfirmation, delight, experiment	

Surprise … Surprise…

An Empirical Investigation on How Surprise is Connected to Customer Satisfaction

JOELLE VANHAMME *

ABSTRACT

This research investigates the specific influence of the emotion of surprise on customer transaction-specific satisfaction. Four empirical studies–two field studies (a diary study and a cross section survey) and two experiments–were conducted. The results show that surprise positively [negatively] influences satisfaction directly and indirectly (via the amplification of positive [negative] emotions), even when disconfirmation is taken into account in the model. The amplification property of surprise and the How-do-I-feel-about-it? heuristic are believed to explain this influence. Some results also show that surprised customers display higher levels of satisfaction and dissatisfaction than non surprised customers.

INTRODUCTION

It has been argued that "businesses need to move beyond mere satisfaction, to customer delight" (Rust, Zahorik and Keiningham 1996, p. 229). Customer delight–which is defined as the highest level of customer satisfaction–is believed to translate into higher customer retention and loyalty (Oliver, Rust and Varki 1997; Rust et al. 1996). And, interestingly, PLEASANT SURPRISE is believed to be a privileged way to trigger customer delight (Rust and Oliver 2000). However, theoretical and empirical investigations in that area remain scarce.

The idea that surprise and delight are related was first suggested by the empirical work of Plutchik (1980) on emotions. His study shows that the word 'delight' represents a mixture of surprise and joy. Westbrook and Oliver (1991; Oliver and Westbrook 1993) also report some indirect evidence for a link between surprise and satisfaction. Using cluster analysis on the emotions consumers experienced during products/services consumption, their studies brought to light a cluster of consumers with high scores of surprise and joy; these consumers were found to be more satisfied than the consumers from any other group. The exploratory surveys by Oliver, Rust and Varki (1997) and Mano and Oliver (1993) also suggest the existence of a relationship between surprise and satisfaction. Their LISREL analysis supported the causal path 'arousal $\Rightarrow$ positive affective reactions $\Rightarrow$ satisfaction.' However, a closer look at their measure for arousal reveals that it captured something rather closer to surprise (the measure was composed of two items rating how "surprised" and how "astonished" respondents were, i.e., two out of the three items of the Izard (1977)'s Differential Emotion Scale (DES) for surprise).

The combination of surprise with negative emotions has also been granted some attention in the literature. Plutchick (1980)'s work shows that mixtures of surprise and anger and of surprise and sadness are best tapped by the words 'outrage' and 'disappointment', respectively. Interestingly, Schneider and Bowen (1999) have conceptualized OUTRAGE as the mirror image of customer

delight (i.e., the highest dissatisfaction level) and believe that outrage leads to defection and 'terrorism' (i.e., the outraged customer tell his story to other customers and exaggerates it with each retelling). In a similar vein, the study by Mano and Oliver mentioned above also showed the causal path 'arousal (same measure as above) $\Rightarrow$ negative affective reactions $\Rightarrow$ dissatisfaction.'

Thus, the empirical studies mentioned above suggest a potential influence–positive or negative depending on the valence of subsequent affective reactions–of surprise on consumer satisfaction.[1] However, these studies did not specifically investigate the surprise-satisfaction relationship in a marketing context (e.g., Plutchik 1980) and/or lack proper measures for the concept of surprise (e.g., Mano and Oliver (1993)). Furthermore, none of these studies experimentally manipulated surprise. The purpose of this article is therefore to contribute in these areas. It is also worth noting that the studies mentioned above did not–as such–provide a conceptual framework for understanding the processes through which surprise could influence satisfaction. Vanhamme and Snelders (2001)'s theoretical work tried to fill this gap while emphasizing the need for empirical evidence with respect to the existence of a relationship between surprise and satisfaction. This article–notably building on this theoretical work–thus aims at addressing this empirical need.

THE EMOTION OF SURPRISE

Some authors have complained about the lack of a theoretical framework for the emotion of surprise in the marketing literature (Oliver, Rust and Varki 1997; Westbrook and Oliver 1991) and have requested that surprise be investigated in a marketing context (Derbaix and Pham 1989). Most knowledge about this emotion is to be found in the literature in psychology. As, for example, Charlesworth (1969), Ekman and Friesen (1975), Izard (1977) and Plutchik (1980), most recent

---

[1] The present article endorses the traditional uni-dimensional view of satisfaction, which posits that 'satisfaction' and 'dissatisfaction' are opposite elements on a continuum (Evrard 1993; Oliver 1997). The word 'satisfaction' will thus be

studies carried out on surprise (e.g., Reisenzein 2000, Schützwohl 1998) consider it to be a NEUTRAL emotion. This relatively short-lived emotion is elicited by a *schema discrepancy* (for example, stimuli that are unexpected or misexpected; Ekman and Friesen 1975) (Scherer 1984; Schützwohl 1998). A schema is a type of private, normally informal, inarticulate, unreflective theory about the nature of objects, events or situations (Rumelhart 1984). Individuals continuously check whether their schemas match the inputs coming from their environment. As soon as inputs diverge from the schema, surprise is elicited.

Schema-discrepancy check is the first stimulus evaluation check (SEC) defined by Scherer (1984). This author assumes that each emotional state is the result of one or more sequential SEC. More specifically, he describes five sequential SEC, the first two being relatively unconscious and automatic: 1) schema discrepancy; 2) intrinsic pleasantness or unpleasantness; 3) goal relevance; 4) coping potential; 5) norm compatibility. As a result of the evaluation of the pleasantness / unpleasantness of the experience–which comes straight after the first SEC–, surprise is often followed by another positive or negative emotion. This gives the impression that surprise is colored either positively (e.g., surprise + joy) or negatively (e.g., surprise + anger) (Ekman and Friesen 1975) and explains why people talk about "good or pleasant surprise" and "bad or unpleasant surprise". If no positive or negative affective reaction follows surprise (i.e., none of the SECs elicited a valenced affective reaction), surprise remains uncolored. In the rest of this article, 'positive surprise' and 'negative surprise' will be used to designate blends of surprise with, respectively, positive and negative affective reactions.

The emotion of surprise is described as a 'syndrome of reactions', that is, a specific pattern of reactions at the SUBJECTIVE (e.g., subjective feeling, surprise exclamation), PHYSIOLOGICAL (e.g., changes in the respiration and heart rates) and BEHAVIORAL levels (e.g., specific facial expression–i.e., raised eyebrows and opened eyes and mouth, interruption of ongoing activities,

---

used to describe both the construct and the positive side of the continuum and 'dissatisfaction' will refer to the negative side of the continuum.

focusing of attention on the surprising stimulus) (e.g., Meyer, Reisenzein and Schützwhol 1997; Vanhamme 2000, see figure 1). Even though the traditional and widespread *syndrome of reactions* concept advocates a strong association between the emotion components (such as facial expressions, subjective experience, physiological changes), several studies have shown that the different components of emotions are only weakly associated or even not significantly associated (Reisenzein 2000). This is also the case for surprise. For example, Niepel (1996, referenced by Reisenzein 2000) obtained significant but weak correlations between facial expressions, psycho-galvanic response and reaction times; Reisenzein (2000), Reisenzein and Bördgen (1998), and Vanhamme (2000) reported strong dissociation between 1) facial expression and subjective experience, 2) sub-components of the facial expression (most often only one sub-component appeared: raised eyebrows) and 3) facial expression and beliefs respondents had about their own expressiveness. The weak correlations between different measures of surprise (e.g., coding of facial expressions, verbally reported surprise) could result from the fact that they do not necessarily reflect the same 'facets' of this emotion and advocate for a looser version of the syndrome of reactions concept (Reisenzein 2000).

Overall surprise results in processes aiming at eliminating the schema discrepancy, that is, the analysis and, if NECESSARY, the updating of the relevant schema (e.g., Meyer et al. 1997).

Insert figure 1 about here

HYPOTHESES: THE LINK BETWEEN SURPRISE AND SATISFACTION

Prior to describing the research hypotheses, it should be emphasized that transaction-specific satisfaction (hereafter referred to as 'satisfaction') has been adopted for this study–like most studies investigating the antecedents of satisfaction (Anderson and Fornell 1994; Oliver 1997). Along with recent definitions (e.g., Aurier and Evrard 1998, Vanhamme and Snelders 2001), this concept is defined as a psychological (i.e., cognitive and affective) state, which results from a buying and/or consumption experience (i.e., transaction specific). This definition clearly emphasizes the dual

character of satisfaction–i.e., it is neither a pure emotion, nor a pure cognition (e.g., Mano and Oliver 1993; Oliver 1993, 1997; Price, Arnould and Deibler 1995; Westbrook 1987; Westbrook and Oliver 1991)–that has been largely acknowledged. Despite this common dual characteristic shared with the concept of attitude, satisfaction is not an attitude but an antecedent of it (Oliver 1981, 1997). And, since "satisfaction soon decays into (but nevertheless greatly affects) one's overall attitude toward purchasing products" (Oliver 1981, p. 51), a measurement of satisfaction immediately after consumption / purchase should provide the best construct validity (Labarbera and Mazursky 1983).

Research on satisfaction dates back to the 1970's and several antecedents of this concept have been brought to light since then. For example, the three components of the Oliver's so called expectation-disconfirmation paradigm (i.e., expectations, performances and disconfirmation) (e.g., Churchill and Surprenant 1982; Oliver and de Sarbo 1988); attributions (e.g., Bitner 1990); equity (e.g., Oliver and de Sarbo 1988); quality (e.g., Oliver 1994); positive and negative affective reactions (e.g., Westbrook 1987). However, these studies have not analyzed the specific role of the emotion of surprise on satisfaction.

Two complementary theoretical explanations can account for the influence of surprise on satisfaction. A first explanation relies on surprise's ability to heighten the intensity of affective reactions that follow it (Charlesworth 1969, Desai 1939, Mellers, Schwarz and Ritov 1999). Vanhamme and Snelders (2001) call this a RESPONSE CONTAGION effect. For example, Mellers, Schwarz and Ritov (1999) show that the pleasure of winning and the displeasure of loosing are more intense when the result of the game is surprising. In other words, affective responses to a stimulus are more extreme when this stimulus elicits surprise. This enhancing/amplification property of surprise is explained by its impact on the organism such as the state of suspended action, physiological activity and change in the subject-object relationship (e.g., heightened consciousness of the surprising object) (Desai 1939). In the context of a consumption / purchase experience, surprise would thus enhance all [negative] positive emotions, such as joy [anger], that follow it. These more intense positive [negative] emotions would then increase [reduce] the satisfaction level of consumers (Westbrook

1987). This path describes the indirect influence of surprise on satisfaction (see path a in figure 2).

However, surprise experienced during a consumption / purchase should also directly lead to a more

extreme satisfaction evaluation since satisfaction is partly affective (= direct influence of surprise on

satisfaction, path b in figure 2).

Insert figure 2 about here

Another complementary explanation for the influence of surprise on satisfaction is based on

the *affect-as-information view* (Schwarz 1990)–and more specifically the *how–do–I–feel–about–it*?

heuristic (Pham 1998; Schwarz and Clore 1996). This heuristic posits that–in cases where people

perceive their affective reactions (e.g., emotion, mood) as relevant to a judgment (e.g., attitude toward

the ad, liking, satisfaction; see Pham et al. 2001)–they use these reactions as a source of information

for making their judgment with an object (Pham 1998; Schwarz and Clore 1996): positive affective

reactions lead to a favorable evaluation and negative affective reactions to an unfavorable evaluation.

Moreover, according to the heuristic, the physiological intensity of the affective reactions experienced

influences the strength of the judgment: consumers–inspecting "how they feel"–infer from their

physiological arousal that they "feel strongly" about the object and, as a result, give more extreme,

more polarized evaluations (Gorn, Pham, and Sin 2001; Pham et al. 2001). Since surprise is

characterized by a state of heightened physiological arousal, a consumer's affective responses are

likely to be more extreme, more polarized if s/he has recently been surprised. This effect operates

through the Zillmann's (1983) excitation transfer hypothesis (Gorn et al. 2001). This theory posits that

physiological arousal decays relatively slowly and that residues of arousal from prior stimulation

combine with excitation in subsequent stimulation; the combined activity is then expected to intensify

the emotional experience during the subsequent stimulation but only if the individual does not

distinguish between portions of his/her excitation that are due to prior excitation and the present

stimulation (e.g., Zillmann 1983). This effect has been reported in a large number of studies (see

Reisenzein 1983 for a review). Most studies (e.g., Cantor, Zillmann and Bryant 1975) induced

physiological arousal through physical exercises (bicycling), which involves clear and readily

distinguishable cues of the source of arousal (e.g., heart pounding, heavy breathing). In those studies

the effect of excitation was only obtained when subjects were exposed to the subsequent stimuli some

minutes after the physical exercise (so that they would not perceive their arousal anymore). It is

important to note, however, that subjects were specifically asked to rate their perceived arousal in

these studies. Therefore, artificially high levels of attention to introspective cues–such as heart

pounding–made subjects more sensitive to their physiological arousal than they would have been

under ordinary conditions. Under more natural conditions–such as during a consumption / purchase

experience–, people are not likely to pay so much attention to the source of their physiological

arousal, which makes excitation transfer easier to occur. Furthermore, the studies by Gorn et al.

(2001) further suggest that the excitation transfer effect is more likely to occur under moderate

physiological arousal because its actual source is likely to be less salient than the source of high

physiological arousal (such as when pedaling). It seems fairly reasonable to argue that the intensity of

physiological arousal experienced during surprising consumption or purchase is usually moderate at

the most. Consequently, excitation transfer is highly likely to occur during surprising consumption /

purchase experiences. Thus according to the *how–do–I–feel–about–it*? heuristic, satisfaction

evaluations are likely to be more extreme, more polarized (i.e., higher [lower]) when the consumption

/ purchase experiences are surprising than when they are not surprising. This is consistent with the

direct effect of surprise on satisfaction represented in figure 2, path b. However, the indirect effect

(figure 2, path a) could also occur through the accentuation/polarization of the positive [negative]

emotions (that accompany surprise) themselves. Hence:

H1a: For positively valenced[2] consumption / purchase experiences, surprise has a positive

direct and indirect influence–through the amplification of subsequent positive

---

[2] 'Positively valenced' and 'negatively valenced' refer to the valence of emotions (other than surprise) experienced

during the consumption / purchase, and resulting from SECs that are subsequent to the schema-discrepancy check.

emotions–on satisfaction. Positive emotions will thus not fully mediate the influence of surprise on satisfaction.

H1b: For negatively-valenced[3] consumption / purchase experiences, surprise has a negative direct and indirect influence–through the amplification of subsequent negative emotions–on satisfaction. Negative emotions will thus not fully mediate the influence of surprise on satisfaction.

Note that Vanhamme and Snelders (2001) suggest an additional explanation that is consistent with hypotheses 1a and 1b, which relies on an *accessibility effect*. The more intense an emotion experienced during the consumption/purchase, the more salient and accessible its memory trace and the higher the likelihood that it will be integrated in the satisfaction judgement (Cohen and Areni 1991; Westbrook and Oliver 1991). Since surprise leads to an interruption of ongoing activities, a focus of the attention, and a heightened consciousness of the surprising stimulus, surprising events are found to leave stronger traces in memory (e.g. Meyer et al. 1997). In addition, the amplification property of surprise may lead to positive or negative emotions that are more intensely experienced during the surprising C/P and this can also lead to more salient and more accessible memory traces. Furthermore, retrieved emotional episodes have the capacity to elicit the same emotions–although not necessarily with the same intensity (Cohen and Areni 1991)–and may thus–in addition–influence satisfaction through a response contagion effect.[3]

Besides the two first hypotheses dealing with the way surprise influences satisfaction, two additional hypotheses can be written concerning its differential impact on the level of satisfaction. Oliver (1997) suggests that the higher the level of [dis]satisfaction, the higher the intrinsic arousal contained in the [dis]satisfaction response. Therefore, due to the arousal that is part of the emotion of

---

[3] Note that Vanhamme and Snelders (2001) make a distinction between memory-based, inference-based and on-line satisfaction judgements. According to these authors, response contagion is the only process through which surprise is likely to influence satisfaction if the satisfaction judgement is formed on-line whereas both response contagion and accessibility effect are likely to be found for memory-based and inference-based satisfaction judgements.

surprise, a positively [negatively] surprised customer should experience a level of [dis]satisfaction with a higher degree of arousal–i.e., a higher level of [dis]satisfaction–than non-surprised consumers. Hence:

H2a: Consumers who are positively surprised during their consumption / purchase experience will have a higher level of satisfaction than if they had experienced the same consumption / purchase without being surprised.

H2b: Consumers who are negatively surprised during their consumption / purchase experience will have a lower level of satisfaction (i.e., a higher level of dissatisfaction) than if they had experienced the same consumption / purchase without being surprised.

The last hypothesis deals with the specific role of surprise compared to *disconfirmation*–a key determinant of satisfaction. Disconfirmation is a COGNITION that summarizes the recognition that the performance of a product/service is better than expected, as expected or worse than expected (Oliver 1997; Westbrook 1987). Among the antecedents of satisfaction, this cognitive variable usually accounts for the largest part of its variance (Churchill and Surprenant 1982; Oliver 1997). Some studies have, however, also shown that positive and negative emotions considerably add to the explanatory power of the satisfaction model (e.g., Oliver 1993), and that the relationship between positive and negative emotions and satisfaction–which is not a purely cognitive variable–is not fully mediated by disconfirmation (Westbrook 1987; Philips and Baumgartner 2002). Surprise being an emotion, disconfirmation is thus also unlikely to encompass ALL ASPECTS of surprise that may influence satisfaction. Hence:

H3: For surprising experiences, the influence of surprise on satisfaction will remain significant once the influence of disconfirmation is taken into account.

GENERAL METHODOLOGY

The hypotheses have been investigated through four different studies: two field studies using two different data collection methods–a diary survey (study 1) and a cross-section questionnaire survey (study 2)–and two laboratory experiments (studies 3 and 4). On the one hand, the field studies allow the study of a wide variety of surprising consumption / purchase experiences and provide more insight into the characteristics of these types of experiences. On the other hand, the laboratory experiments–each based on a single product/service–allow a controlled manipulation of surprise during the consumption / purchase experience and the use of non-verbal measures of emotions beyond questionnaire measures (i.e., verbal reports of emotions). The marketing literature has traditionally favored verbal reports of emotions (e.g., Richins 1997) but–besides the fact that no single measure is able to record all facets of an emotion (Derbaix and Poncin 1999)–this type of measure displays numerous drawbacks. Participants tend to rationalize their emotions (one way to reduce this is to measure the emotions immediately after they have been elicited); they cannot always identify their emotional state (because they lack abilities of introspection and retrospection); and they may not want to share their emotions with strangers (Derbaix and Poncin 1999). Some researchers have, therefore, sought to develop non-verbal measures of emotions (e.g., physiological measures, coding of facial expressions). The results of her study lead Vanhamme (2000) to advise the use of both verbal reports and coding of facial expressions in order to measure surprise in a marketing context.[4] Although these measures have advantages (they record emotions in real time and are free of retrospection and introspection biases), they can also be costly, time consuming and constraining. For

---

[4] In her experiment, she used and compared verbal report (Izard's DES), galvanic skin response (GSR), electromyography (EMG) and coding of facial expressions of respondents. It was shown that GSR and EMG measures were not well suited for consumption / purchase settings. Consumers do not usually stand stock-still while 'consuming' the P/S, which elicits changes in GSR and EMG that are not related to emotions experienced. Coding of facial expressions gave the best results (and significantly explained satisfaction scores).

example, it can be very difficult to record facial expressions with enough precision outside of a laboratory, in a natural setting (this explains why these facial measures were only used in studies 3 and 4). Note that hypotheses 2a and 2b will be investigated in the experimental studies (studies 3 and 4). A preliminary analysis of these hypotheses will, however, also be carried out in study 2.

STUDY 1

The purpose of study 1 was to test hypotheses 1a, 1b, and 3 using a field study–more specifically a diary survey. The diary technique aims at gathering information about particular events, but this gathering is done over an extended period of time, as well as at pre-specified moments. To a large extent, this method avoids problems of retrospective reporting from memory but may increase reactivity from the respondents since they know the type of events they will be asked to describe and might, therefore, pay more attention to these events than what they otherwise would have done. The diary technique is widely used in psychology (Stone, Kessler, and Haythornthwaite 1991; Wheeler and Reis 1991) and has also been used in consumer behavior (e.g., the study by Price, Arnould, and Deibler (1995) of emotional responses to service encounters).

Method

Prior to the final data collection, a pre-test of the questions included in the diary was carried out on 18 respondents. The purpose was to make sure that all questions were properly understood.

Twenty-seven respondents participated in the diary study (18 women, nine men, all middle-class with an age range of 27 to 70 years and an average age of 41 years, all non-students). They were asked–in the first part of the diary–to write down (each night over a seven-day period) all their purchase experiences–followed or not by consumption–(referred to as 'consumption/purchase experiences hereafter) and specify whether they were surprising or not. In the second part of the diary, the respondents were asked to fully describe their last surprising experience and their last non-surprising experience for that particular day and, then, to answer questions (i.e., emotions,

satisfaction, and disconfirmation scales) about them.[5] The scales used were the multi-item DES (Izard 1977) for measuring surprise and the other emotions (3 items per emotion); a mono-item scale of disconfirmation (Oliver 1997), and the mono-item DT scale (Westbrook 1987) (see appendix A). Westbrook (1980) has shown the reliability of the DT scale and its convergence with other satisfaction scales. The DT scale was chosen for theoretical reasons. First, it allows for explicit reference to the affective dimension of satisfaction and a better grasp of this dimension compared to traditional satisfaction scales. Second, it also provides for a better differentiation of the positive extreme of the satisfaction continuum (Westbrook 1980). However, since the DT scale is a mono-item scale, respondents were also asked to answer the multi-item Oliver's satisfaction scale [SAT] (Oliver 1980).[6] All 27 respondents participated in a training session organized the night before the beginning of the data collection.

Results

Usual checks for unidimensionality (using Principal Component Analysis (PCA)), reliability and normality were carried out for each measure. All Cronbach *alpha* for multi-items measures collected under the study were above .7, except for fear (.65). Since the assumption of normality of the distribution did not hold for most of the variables, non-parametrical 'distribution-free' analyses will be carried out. Hypotheses 1a, 1b and 3 will thus be tested using Kendall T and Kendall partial T ranks-order correlation coefficients. These coefficients are considered to be robust alternatives for

---

[5] If no surprising [non surprising] consumption / purchase had been experienced that day, respondents had to fill in the second part of the diary for their last two non surprising [surprising] consumption / purchase experiences.

[6] Some control variables were also collected (gender, age, personality traits (Saucier 1994)). However, these variables did not influence the pattern of results and will thus not be mentioned further.

Pearson correlation and Pearson partial correlation coefficients (Choi and Marden 1998; Rao and Sievers 1995).[7]

As could be expected, respondents reported (in the first question of the diary) on average more non-surprising consumption / purchase experiences than surprising experiences (12.61 versus 4.58 per week and per respondent respectively, Wilcoxon signed ranks test Z = -4.147, $p_{1\text{-}tail} < .001$). The surprising experiences involved products (e.g., food, toys, durable goods, clothing, cleaning, and hygiene products) in about 60% of the cases and services (e.g., restaurants, cable TV service, post office, doctors, hairdressers, information services) in about 40% of the cases. The reported causes of surprising experiences were related to newness of the product, quality-to-price ratio, quality, choice, out-of-stock problems, product or service failure, coincidence, gifts of the seller, factors linked to atmospherics, and differences between what was ordered and what was delivered.

In total, detailed descriptions of 265 consumption / purchase experiences were collected in the second part of the diary. Among them, 174 were 'POSITIVELY-VALENCED EXPERIENCES' (i.e., experiences for which respondent's emotions (other than surprise) were mainly positive; 104 of the positively-valenced experiences were surprising [POSITIVELY SURPRISING EXPERIENCES]); 50 were 'NEGATIVELY-VALENCED EXPERIENCES' (i.e., experiences for which respondent's emotions (other than surprise) were mainly negative; 37 of the negatively-valenced experiences were surprising [NEGATIVELY SURPRISING EXPERIENCES]); and 41 were 'NEUTRAL EXPERIENCES' (of these 90% were experiences free of positive and negative emotions and 10% were experiences that elicited positive emotions and negative emotions with the same low intensity; eight of the neutral experiences were surprising). Note that, as could be expected for consumption / purchase experiences, there was a larger proportion of positively-valenced experiences than negatively-valenced experiences (Fiske 1980).

---

[7] Note that the data collection method could have introduced some interdependency in the data since the same respondent provided several observations. This could have affected the results of the statistical tests. However, bootstrap analyses were carried out on the data and showed that this was not the case.

As expected for positively-valenced consumption / purchase experiences (hypothesis 1a), the results show a direct and indirect (through positive emotions) link between surprise and satisfaction. Surprise is positively correlated with joy (Kendall $T = .250$, $p_{1\text{-}tail} < .01$; N = 174) and the latter is positively correlated with satisfaction (DT : Kendall $T = .294$, $p_{1\text{-}tail} < .01$; N = 170 // SAT : Kendall $T = .207$, $p_{1\text{-}tail} < .01$; N = 165). Surprise is also positively correlated with satisfaction–but only for the DT scale (Kendall $T = .301$, $p_{1\text{-}tail} < .01$; N = 170)–and this correlation remains significant once 'joy' is controlled for (Kendall $T_{\text{Sur-DT . Joy}} = .246$, $p_{1\text{-}tail} < .001$).[8]

As shown in table 1, hypothesis 1b is also supported by the data collected for negatively-valenced experiences. Surprise is positively correlated with all negative emotions (except for sadness) and the latter are negatively correlated with satisfaction. Surprise is also negatively correlated with satisfaction (table 1, section A). And, this correlation remains significant once negative emotions are controlled for (table 1, section B).

<div style="text-align:center">Insert table 1 about here</div>

Furthermore–as expected (hypothesis 3) for positively and negatively surprising experiences–the significant correlations that exist between surprise and satisfaction remain significant once controlled for disconfirmation (POSITIVELY SURPRISING EXPERIENCES: Kendall $T_{\text{Sur-DT.Disc}} = .202$, $p_{1\text{-}tail} < .01$; N = 101; Kendall $T_{\text{Sur-SAT.Disc}} = .169$; $p_{1\text{-}tail} < .01$; N = 100 // NEGATIVELY SURPRISING EXPERIENCES: Kendall $T_{\text{Sur-DT.Disc}} = -.209$; $p_{1\text{-}tail} < .05$; N = 35). The partial correlation with SAT, controlling for disconfirmation, is however not significant in the sample of negatively surprising experiences

---

[8] Note that for positively surprising experiences, the correlation and partial correlation between surprise and SAT become significant (Kendall $T_{\text{Sur-SAT}} = .240$, $p_{1\text{-}tail} < .01$; N = 100; Kendall $T_{\text{Sur-SAT . Joy}} = .15$, $p_{1\text{-}tail} < .05$) (all the other correlations between variables also remain significant for this 'positively surprising' sample and are even stronger).

Conclusion and Discussion

The data supported the hypotheses. A direct and an indirect link–through positive emotions–was found between surprise and satisfaction for positively-valenced experiences. And, similarly for negatively-valenced experiences, data showed a direct and an indirect link–through negative emotions–between surprise and satisfaction. In both cases, the direct link between surprise and satisfaction was thus not totally mediated. The data also confirmed that the influence of surprise on satisfaction is not totally taken into account by disconfirmation and that surprise thus deserves to be measured on its own. In addition, these results clearly emphasize–like prior research (e.g., Westbrook 1987)–the need to further study and incorporate emotions in satisfaction models and not just cognitive variables.

The limitations of this study need however to be acknowledged. The sample was a convenient sample; the measures of emotions were purely and only verbal (introspection bias might thus be present in the results); the analyses are only correlational and do not allow conclusions to be drawn in terms of causality (however, it seems very unlikely that satisfaction would cause surprise). Furthermore, the rather large amount of surprising experiences that was collected could partially be due to the fact that the data collection technique made respondents more aware of their emotions during their consumption / purchase experiences. This is a limitation of the technique (Wheeler and Reis 1991) since it allows respondents to anticipate and pay more attention to the events they will have to report and can thus lead to a greater reactivity. However, since none of the respondents guessed the goal of the study, it is not likely that the potential reactivity changed the relationship between surprise and satisfaction. Finally, retrospection bias remains limited compared to other questionnaire surveys (Wheeler and Reis 1991) since diaries were filled in very soon after the experience.

STUDY 2

The purpose of Study 2 was to replicate the results of the diary study with respect to hypotheses 1a, 1b, and 3 using another type of field study–a cross section questionnaire survey designed to specifically tap each respondent's most recent surprising consumption/purchase experience–, which avoids potential reactivity biases. Study 2 additionally aimed at providing a preliminary analysis of hypotheses 2a and 2b.

Method

Prior to starting the final data collection for study 2, questionnaires were pre-tested on 24 respondents to ensure proper understanding of the questions.

Participants were 100 non-students respondents drawn from within 30 to 50-year-old consumers coming from the middle and upper-middle classes. Half of them at random (hereafter referred to as 'subsample A') were first instructed to recall and explain in detail their most recent consumption / purchase experience (with a product or a service) which surprised them positively (this was part 1 of the questionnaire). Next, they were asked to answer the same questions for their last experience with the same kind of product/service that did not surprise them (this was part 2 of the questionnaire). The other half of the participants (hereafter referred to as 'subsample B') was assigned the same task except that part 1 of the questionnaire concerned their most recent negatively surprising experience. Thus, the total sample includes 100 surprising experiences and 100 non-surprising experiences. The measures were the same as for study 1: the multi-item DES for emotions, the mono-item disconfirmation scale, as well as the DT and SAT satisfaction scales.[6]

Results

Usual checks for unidimensionality (using PCA), reliability and normality were carried out for each measure. All Cronbach *alpha* for multi-items measures collected under the study are above .7.

Since the assumption of normality of the distribution did not hold for most of the variables, hypotheses 1a, 1b and 3 will be tested using Kendall T and Kendall partial T ranks-order correlation coefficients–as for study 1–and hypotheses 2a and 2b will be tested using the Wilcoxon signed ranks test Z (Siegel and Castellan 1988).

As with the diary study, surprising experiences involved products (e.g., food, electrical appliances, cosmetics, cars, clothing) in about 60% of the cases and services (e.g., restaurants/bars, movies, information services, theatre) in about 40% of the cases. The same kind of product/service was involved in both positively and negatively surprising experiences. The reported causes of positively and negatively surprising experiences were very similar to the ones reported for the diary study; they were related to newness of the product, quality-to-price ratio, out of stock problems, product / service failure, poor quality, factors linked to atmospherics and differences between what was ordered and what was delivered.

The data seem to support hypotheses 2a and 2b since respondents had higher scores of satisfaction for their positively surprising experience (PoSE) than for their non-surprising experience (NonSE) with the same product/service (Median DT $_{PoSE}$ = 7.0 vs. Median DT $_{NonSE}$ = 5.0; Wilcoxon signed ranks test Z = -6.029, $p_{1\text{-}tail}$ < .001 // Median SAT $_{PoSE}$ = 10.0 vs. Median SAT $_{NonSE}$ = 7.6; Wilcoxon signed ranks test Z = -5.77, $p_{1\text{-}tail}$ < .001) and lower scores for their negatively surprising experiences (NeSE) than for their non-surprising experiences with the same product/service (Median DT $_{NeSE}$ = 2.0 vs. Median DT $_{NonSE}$ = 5.0; Wilcoxon signed ranks test Z = -6.151, $p_{1\text{-}tail}$ < .001 // Median SAT $_{NeSE}$ = 1.0 vs. Median SAT $_{NonSE}$ = 8.3; Wilcoxon signed ranks test Z = -6.140, $p_{1\text{-}tail}$ < .001). These results should, however, be considered with caution as the difference between surprising experiences and non-surprising experiences could possibly be due to elements other than the experience of surprise, such as the perceived performance (even though the type of product/service was the same).

For positively surprising experiences, the results highlight–as expected (hypothesis 1a)–a positive direct and indirect link (through positive emotions) between surprise and satisfaction.

Surprise is positively correlated with joy (Kendall $T = .502$, $p_{1\text{-}tail} < .01$; N = 50) and the latter is positively correlated with satisfaction (DT: Kendall $T = .560$, $p_{1\text{-}tail} < .01$; N = 50 // SAT: Kendall $T = .356$, $p_{1\text{-}tail} < .01$; N = 50). Surprise is also positively correlated with satisfaction (Kendall $T = .415$, $p_{1\text{-}tail} < .01$; N = 50 // SAT: Kendall $T = .391$, $p_{1\text{-}tail} < .01$; N = 50), and this correlation remains significant once 'joy' is controlled for (Kendall $T_{\text{Sur-DT . Joy}} = .19$, $p_{1\text{-}tail} < .05$; Kendall $T_{\text{Sur-SAT . Joy}} = .26$, $p_{1\text{-}tail} < .01$).

For negatively surprising experiences, table 1 shows that surprise is–as expected (hypothesis 1b)–positively correlated with all negative emotions (except for anger) and the latter are also positively correlated with the satisfaction scale. However, negative emotions are not correlated with SAT and surprise is not correlated with DT and SAT (table 1, section C). This could be caused by a lack of variation for some of the variables (which would not leave enough room for significant correlations between the intensity of surprise and the other variables, such as satisfaction).[9] In order to increase variation in the scores, a new 'NEGATIVE SAMPLE' (N = 100) was built. This sample is composed of the answers of subsample B for their negatively surprising experiences and of the answers of subsample A for their non-surprising experiences. In a similar way, a 'POSITIVE SAMPLE' (N = 100) was also built: it is composed of the answers of subsample A for their positively surprising experiences and of the answers of subsample B for their non-surprising experiences. These 'new samples' were so built in order to preserve independence between observations. Before building them, the comparability of both sets of 50 non-surprising experiences was checked. No significant difference was found–between the non-surprising experiences of subsample A and the non-surprising experiences of subsample B–for the variables related to emotions (i.e., DES scale). Similar types of products/services were involved in both subsamples. Furthermore, respondents of subsample A and B were comparable: they had experienced the same kind of products and services, lived in the same

---

[9] For example, scores of surprise for surprising experiences are relatively high and homogenous in the negatively surprising sample. The use of the questionnaires asking for 'extreme experiences' (i.e., positively or negatively surprising experiences) as was done in this study could be responsible for the problem of low variation in the scores.

urban areas, had the same types of leisure activities and were similar in terms of sex, age and personality (using Saucier (1994) mini-markers).

When using the NEGATIVE SAMPLE of 100 observations, a significant correlation appears between surprise and anger, between surprise and satisfaction, and between all negative emotions and satisfaction (all other correlations between variables remain significant but are much stronger) (table 1, section D). Furthermore, the significant correlations between surprise and satisfaction remain significant after controlling for the influence of negative emotions (see table 1, section E). Hypothesis 1b seems thus to be confirmed. Note that for the POSITIVE SAMPLE the results confirm those of the original sample (i.e., the positively surprising sample): correlations are about the same as, or even stronger than the ones in the positively surprising sample (e.g., Kendall $T_{\text{Sur-DT}} = .61$, $p_{\text{1-tail}} < .01$; Kendall $T_{\text{Sur-DT . Joy}} = .47$, $p_{\text{1-tail}} < .01$; N = 100).

As expected (hypothesis 3), the significant correlation that exist between surprise and satisfaction remains significant (for $\alpha = .09$ for the negative sample, though) once controlled for disconfirmation. The correlation with the SAT scale in the *negative sample* is, however, only nearly significant for alpha = .1 (POSITIVELY SURPRISING EXPERIENCES: Kendall $T_{\text{Sur-DT.Disc}} = .21$, $p_{\text{1-tail}} < .05$; Kendall $T_{\text{Sur-SAT.Disc}} = .26$, $p_{\text{1-tail}} < .01$; N = 50 (results are similar for the POSITIVE SAMPLE*) // *NEGATIVE SAMPLE: Kendall $T_{\text{Sur-DT.Disc}} = -.09$, $p_{\text{1-tail}} < .09$; Kendall $T_{\text{Sur-SAT.Disc}} = -.08$, $p_{\text{1-tail}} = .11$; N = 100).

Conclusion and Discussion

The data supported all the hypotheses. A direct and indirect link–through positive or negative emotions–was found between surprise and satisfaction. And the results further showed that the influence of surprise on satisfaction was neither totally mediated by these positive or negative emotions, nor totally captured by disconfirmation. The results related to the negative sample should, however, be considered with caution due to the artifice used for increasing variance.

Further limitations of this study are the convenient sample; the purely verbal measures of emotions (and thus possibly an introspection bias); weak variances for some variables; and correlational analyses (but again, a 'satisfaction =>surprise' relationship is theoretically very unlikely). Although the data collection technique is subject to retrospection bias this latter was limited by two aspects. First, surprising consumption / purchase experiences can be considered as outstanding, extreme experiences since they are not likely to be frequent among all consumption / purchase experiences and, as such, can be more accurately recalled than those which are more mundane (Flanagan 1954). Second, reported experiences were recent and consumers were motivated to engage in detailed observations/evaluations during the occurrence of the surprising incident (surprise is known to lead to an analysis of the event). Eventually, for both studies 1 and 2, a single mono-item measure of disconfirmation was used, which could be seen as a limit with regard of the results for hypothesis 3. To conclude, it should however be emphasized that–despite their respective limitations–the largely similar results reported for those two field studies gives some confidence in their ability to reflect the reality and highlight the need to investigate the specific role of surprise, besides cognitive variables, in satisfaction models.

STUDY 3.

The purpose of study 3 is to replicate results from the two field studies with respect to hypotheses 1a, 1b and 3–using an experimental manipulation of surprise and non verbal measures of emotions–and additionally test hypotheses 2a and 2b. Compared to the field studies, the experimental design allowed a test of the hypotheses in a controlled environment and the use of more sophisticated measures of emotions.

Method

The experimental design was a "after only with control group" design (Churchill, 1999) including two experimental groups (EG^+ & EG^-) and two control groups (CG^+ & CG^-). The products

used during the experiment were strawberry yogurts wrapped in individual packages. Participants–recruited among a non-student population (age between 20 and 55)–were matched according to gender, age and the frequency of consumption of strawberry yogurt, and then randomly assigned to the four conditions. Participants had to perform four tasks during the experiment; the manipulation of surprise occurred–in EG^+ and EG^-–during the THIRD TASK. During this third task participants had to taste three jars of yogurt. In EG^+, surprise was elicited by the discovery of a foldable plastic spoon hidden under the wrapping of the THIRD YOGURT jar. This was supposed to be a pleasant surprise. The participants in EG^+ did not know that they would discover such a foldable plastic spoon in the wrapping of the third jar of yogurt (the two previous yogurts did not contain a foldable spoon and the existence of such a spoon was not mentioned at any time). In EG^-, surprise was elicited by the absence of the foldable spoon in the third yogurt. This was supposed to be an unpleasant surprise since the foldable spoon had been presented as part of the product from the start (and there was a foldable spoon in the two first strawberry yogurts). In CG^+, the foldable spoon was presented as part of the product from the start and there was a foldable spoon in all strawberry yogurts–as announced. In CG^-, there was no mentioning of a foldable spoon and such a spoon never appeared in any of the yogurts. Thus, nothing was surprising for the participants of CG^+ and CG^-. Pre-tests were carried out prior to the experiment in order to check the manipulation and the material.

When each participant arrived in the laboratory, the experimenter showed him/her a demonstration videotape describing and illustrating the four tasks to be performed.[10] Beside the description of the experimental tasks, the demonstration tape's purpose was also to create the participants' schema about the product characteristics (i.e., a strawberry yogurt with the foldable spoon in the CG^+ and EG^- conditions and without it in the two remaining conditions). During the first task, participants were asked to test and rank different plastic spoons by order of preference. To do so,

---

[10] Since respondents knew – prior to performing the tasks – that they would have to evaluate their consumption/purchase experience, the satisfaction judgement was most likely formed on-line and surprise was most likely to affect satisfaction through response contagion (see note 4).

they were given a strawberry yogurt identical to the ones they would have to taste during the third task; the jar was still in its packaging and included a foldable spoon in the CG$^+$ and EG$^-$ conditions. The purpose of this task was, 1) to make sure that the participants would not be surprised–during the third task–by other elements (e.g., the taste of the yogurt) than the discovery of the foldable spoon and, 2) to reinforce the schema of the product. The second task was a questionnaire with items about involvement with the product-category, mood and intro-/extraversion. Then respondents performed task 3, and immediately afterwards answered the final questionnaire including a manipulation check section and scales about emotions, satisfaction, disconfirmation and appreciation of the foldable spoon (task 4). The participants were left alone while performing the tasks. The whole task lasted for a maximum of 45 minutes, including the debriefing.

With respect to the measures, the Izard (1977)' DES was used to capture the subjective experience of surprise and of positive and negative emotions (i.e., multi-items verbal reports on a five points scale of intensity). Facial expressions were video recorded and coded independently by three judges according to the procedure developed by Ekman and Friesen (1975) and adapted–for surprise– by Reisenzein (2000). Other observable aspects pertaining to the emotion of surprise were also coded (i.e., focus of attention, interruption of ongoing activities, surprise exclamation). Beside the verbal and non-verbal measures of emotions, a three items (10 points scale) measure of the appreciation/liking of the foldable spoon as well as a multi-items scale of disconfirmation (Oliver 1997) and three scales of satisfaction were administered (the mono-item DT scale (Westbrook 1980), the Oliver (1997)´s scale (SAT), and the SATEXP scale of Aurier and Evrard (1998)). Control measures of mood (Peterson and Sauber (1983) scale), extraversion/introversion (Saucier (1994) scale) and involvement with the product-category (Jain and Srinivasan 1990) were also collected.

Results

One hundred and thirty valid questionnaires were collected (31 for EG$^+$; 31 for CG$^+$; 34 for CG$^-$; 34 for EG$^-$). Sixteen questionnaires had to be discarded mainly due to either technical problems,

obvious demand effects, or subjects not following the instructions. None of the remaining 65 participants from the two CG declared having experienced surprise whereas all remaining 65 participants in the two EG were surprised (this surprise was pleasant for EG^+ and unpleasant for EG^-: $mean_{EG+}$ 7/10 and $mean_{EG-}$ 3/10 on the scale anchored by "very negative (1)" to "very positive (10)"). All but four participants in EG^+ and one in EG^- displayed at least one observable characteristic of surprise and no participant from the control conditions displayed characteristics of surprise. Highly significant differences were found between the two EG and the two CG for the measures of surprise. Corroborating the work by Reisenzein (2000) and Vanhamme (2000), surprise was most often indicated by raised eyebrows (35% of the respondents in EG^+ and 50% of the respondents in EG^-), then opened eyes (EG^+: 13%; EG^-: 38%); opened mouth happened the least often (EG^+: 6%; EG^-: 26%). As previous studies (e.g., Reisenzein 2000, Vanhamme 2000), the full facial expression of surprise (i.e., eyebrows + eyes + mouth) appeared for only very few respondents within the surprise conditions (7.6%). Interruption of ongoing activities and focus of the attention/exploratory behavior were observed for, respectively, 61% and 68% of the respondents in EG^+ and for, respectively, 41% and 85% of the respondents in EG^-. Finally, a few respondents had a spontaneous surprise exclamation (10% in EG^+ and 38% in EG^-).

Scales were assessed for uni-dimensionality using PCA. Items were then aggregated and the distributions of each variable checked for normality in each group. The normality assumption did not hold for, 1) verbal surprise (VSUR) for CG^-, 2) verbal negative emotions (VNE) for CG^-, CG^+, and EG^+, and 3) all non-verbal measures–except the overall measure of observed surprise (OBSS) for the two EG and the total facial expression of surprise (FS) for EG^-. Parametrical statistics are thus not appropriate for those variables. Reliability analysis was also performed on all multi-items verbal scales. All Cronbach *alpha* were above .7 except for three of the five dimensions of involvement and intro-/extraversion (these measures are control measures). All inter-judges agreements for facial coding and other observed measured were very high (Rust and Cooil (1994) PRL index >.9). As previous studies (e.g., Reisenzein 2000, Vanhamme 2000), non-significant or weak correlations were

found between verbal and non-verbal measures of emotions (table 2). As expected no difference was found between the two EG and the two CG for any control measures and these measures did not influence the pattern of results (they will thus not be mentioned further).

<div style="text-align:center">——————————————
Insert table 2 about here</div>

Contrary to what was expected (hypothesis 2a), none of the scores of satisfaction are significantly higher for EG^+ than for the two CG (table 3, upper part).[11] However, it is worth noting that surprise has had an impact on the evaluation of the element used for eliciting surprise: in EG^+, the liking of the foldable spoon is significantly higher than in CG^+ whereas the respondents received exactly the same foldable spoon in both conditions (table 3, upper part).

<div style="text-align:center">——————————————
Insert table 3 about here</div>

As expected (hypothesis 2b), satisfaction scores are lower in EG^- than in the CG (the difference for SAT is however not significant). And the scores for the liking of the foldable spoon are also lower in EG^-, although respondents in both groups received exactly the same spoon.

Hypotheses 1a, 1b and 3 were tested by means of correlational analysis (between each individual indicators/measures of each construct; see table 2) and the PLS approach for modeling structural equations (with mode A of estimation)[12] (Wold 1980). Fornell and Bookstein (1982) recommend using this PLS approach as an alternative to the maximum likelihood method of

---

[11] However, it should be noted that only 5 participants in EG^+ used the foldable spoon for tasting the third yogurt in task 3. The others used another spoon (often because they had used a spoon with another design for tasting the first two yogurts and thought it was important to continue with the same design). The 5 participants who used the foldable spoon did have higher satisfaction scores than the others in EG^+, except for DT (SAT: 6.7 vs. 5.5; SATEXP: 6.9 vs. 6.0; SATGLO: 6.7 vs. 5.7; DT: 4.6 vs. 4.6. These differences are, however, not statistically significant, most likely due to the sample sizes). These scores are also higher than those of the control groups (see table 3).

estimation (e.g., LISREL) when–as is the case here–its underlying assumptions are not respected (e.g., too small sample size, non multi-normality) (see Fornell and Bookstein (1982) for a comparison of the two methods). Furthermore, mode A of estimation is free of multi-colinearity problems, which is an advantage as the indicators of a latent variable can be highly correlated. Since both analyses lead to the same general conclusions and the PLS approach allows a report of the results more clearly and more synthetically, the results for the correlational analysis will not systematically be reported here.

The model including all possible indicators for each of the latent variables in EG^+ (i.e., joy, surprise, disconfirmation and satisfaction) had a good predictive validity (i.e., significant Q^2 and R^2 for satisfaction, Fornell & Bookstein 1982; Tenenhaus 1998; Valette-Florence 1988) and a good convergent validity for joy (.65>.5) and satisfaction (.83>.5) (Valette-Florence 1988). For surprise, convergent validity was, however, not good enough (.41 < .5). The model also had a good discriminant validity (.36 < .41) (Valette-Florence 1988). Due to the convergent validity problem for surprise, the model was re-estimated using only the indicators of surprise that were sufficiently correlated with the estimates of the latent variable as advised by Wold (1980). Results are provided in figure 3 (section A). This model is very close to the first model and respects all criteria of predictive (significant Q^2 and R^2), convergent (.66 < .5, .65 < .5 and .83 < .5) and discriminant (.32 < .65) validity.[13]

---

[12] As indicators were reflective, mode A of estimation was used. Note also that Tenenhaus (1998) advises using mode A systematically because of multi-colinearity problems occurring with mode B. Finally, Dijkstra (1981) also recommends mode A when there are a lot of indicators compared to the number of observations (which is the case in this study).

[13] It should be noted that verbal surprise does not belong to the final model because it was not correlated with non verbal measures of surprise. However, correlational results for verbal surprise with respect to hypotheses 1a and 3 were convergent with the PLS model. Verbal surprise was positively correlated with satisfaction and verbal joy. The latter was not significantly correlated with satisfaction and did not mediate the influence of verbal surprise on satisfaction (see table 2). Finally, verbal surprise also remained significantly correlated with satisfaction - but only for the DT scale - once the influence of disconfirmation was controlled for (Pearson $r_{\text{VSUR-DT.DISC}} = .294$; $p_{1\text{-tail}} < .07$; Kendall $T_{\text{VSUR-DT.DISC}} = .275$, $p_{1\text{-tail}} < .05$).

The analysis of the PLS model shows that hypothesis 1a is only partially supported by the data: there seems to be only a direct influence of surprise on satisfaction. Surprise does amplify satisfaction and joy since path coefficients [and correlations] between the latent variables of surprise and satisfaction and between the latent variables of surprise and joy are significant. But the latent variable of joy has no impact on satisfaction (the path coefficient and the correlation between the latent variables of joy and satisfaction are not significant). With respect to hypothesis 3, the direct influence of surprise on satisfaction remains–as expected–significant despite the presence of disconfirmation in the model.

Insert figure 3 about here

The model including all possible indicators for each of the latent variables in EG⁻ (i.e., indicators of negative emotions, surprise, disconfirmation and satisfaction) also suffered from a problem of convergent validity for surprise. The predictive validity and convergent validity for satisfaction were good (i.e., significant Q^2 and R^2 for satisfaction; .76>.5). The model has thus been re-estimated using only the indicators of surprise that were sufficiently correlated with the estimates of the latent variable. Results are provided in figure 2 (section B). This model respects all criteria of predictive (significant Q^2 and R^2), convergent (.71 < .5 and .76 < .5) and discriminant (.59 < .71) validity. It is very close to the first model except that the path and correlation coefficients between the latent variables of surprise and negative emotions (and their significance level) drop heavily (from .490 ($p_{1\text{-}tail} < .01$) to .213 ($p_{1\text{-}tail} = .11$)). This is due to the fact that the new model does not include verbal surprise as an indicator. This variable strongly correlated with verbal negative emotions, which was the single indicator for negative emotions (the judges did not observe any facial expression of negative emotions).[14]

---

[14] As shown in table 2, correlational results for verbal surprise with respect to hypotheses 1b and 3 were convergent with the PLS model. Verbal surprise was negatively correlated with the DT scale of satisfaction (non-parametrical tests also indicated a significant correlation ($\alpha = .1$) for SAT) and positively correlated with verbal negative emotions (see

It should be noted that the correlation between the latent variable of surprise and the verbal indicator of surprise was positive–in the full model with all indicators of surprise–whereas correlations between the latent variable of surprise and non-verbal indicators of surprise were positive. Since verbal surprise is not included in the re-estimated model, all indicators of surprise are negatively correlated with the latent variable. This can be explained by the use of so-called 'display rules' which refers to the tendency to hide or modify one's facial or body expression in public (Ekman and Friesen (1975)).[15]

The analysis of this model shows that hypotheses 1b and 3 are supported by the data (see figure 2, section B). Surprise does amplify satisfaction and negative emotions since path coefficients between the latent variables of surprise and satisfaction and between the latent variables of surprise and negative emotions are significant (for $\alpha = .11$, though). And negative emotions have a significant impact on satisfaction. Furthermore, as expected (hypothesis 3), the influence of surprise on satisfaction remains significant despite the presence of disconfirmation in the model.

Conclusion and Discussion

Within the positive surprise condition, variation in the level of surprise (elicited by the foldable spoon) was found to have a positive DIRECT effect on satisfaction. The direct effect was the sole effect since surprise did not influence satisfaction indirectly (through its amplification of the

---

table 2); negative emotions did not fully mediate the surprise-satisfaction relationship (VNE = f(VSUR) $\Rightarrow P_{1\text{-}tail}$ = .005; DT = f(VSUR) $\Rightarrow P_{1\text{-}tail}$ = .038; DT = f(VSUR, VNE) $\Rightarrow$ **VSUR: $P_{1\text{-}tail}$ = .10,** VNE: $P_{1\text{-}tail}$ = $n.s.$; Pearson $r_{\text{VSUR-DT.VNE}}$ = -.226, $P_{1\text{-}tail} < .10$; Kendall $T_{\text{VSUR-DT.VNE}}$ = -.252, $P_{1\text{-}tail} < .05$) and the latter remained significant once the influence of disconfirmation was controlled for (Pearson $r_{\text{VSUR-DT.DISC}}$ = -.272, $P_{1\text{-}tail} < .07$; Kendall $T_{\text{VSUR-DT.DISC}}$ = -.244, $P_{1\text{-}tail} < .05$).

[15] The pattern of correlations between verbal and non-verbal indicators of emotions for the negative surprise condition also reflects this phenomenon: correlations between verbal and non-verbal measures of emotions are negative whereas they are positive for the positive surprise condition (e.g., GE+: T Kendall $_{\text{JOY-FJ}}$ = .284, $p_{1\text{-}tail} < .05$; Pearson $r_{\text{JOY-FJ}}$ = .309, $p_{1\text{-}tail} < .05$; N = 31 versus GE-:T Kendall $_{\text{JOY-FJ}}$ = -.270, $p_{1\text{-}tail} < .05$; Pearson $r_{\text{JOY-FJ}}$ = -.320, $p_{1\text{-}tail}. < .05$; N = 34). Furthermore, the judges did not observe any facial expression of negative emotions.

emotion of joy). The results also showed that pleasantly surprised consumers did not report significantly higher satisfaction scores than non-surprised consumers. However, this first experiment clearly showed that surprise has had an effect on the evaluation of the surprising product attribute: consumers evaluated more positively the SAME product attribute (i.e., the foldable spoon) in the positive surprise condition than in the non-surprise condition.

The fact that significantly higher scores emerged for the evaluation of the foldable spoon (the surprising product attribute itself) and not for the scores of satisfaction with the whole product could be explained by the relative relevance of the emotion of surprise for the judgement. Participants might have considered their emotion of surprise less relevant for the evaluation of their satisfaction with the whole consumption experience than for the evaluation of the spoon and, as a result, granted this emotion a place large enough to make a significant score increase only in the case of the spoon evaluation. This explanation is consistent with the *How-do-I-feel-about-it?* heuristic. Note, however, that small variances (especially for emotions) cannot be ruled out as a potential alternative explanation for the non-significant results.

Within the negative surprise condition, both a negative direct and indirect (through negative emotions) relationship between surprise and satisfaction was found. Furthermore, respondents in the negative surprise condition had lower scores of satisfaction and worse evaluation of the foldable spoon than the respondents in the non-surprise conditions.

Finally, disconfirmation did not capture the influence of surprise on satisfaction in any of the surprising conditions (i.e., EG^- and EG^+).

As far as the limitations are concerned, some issues related to external validity have to be raised. Even though the non-student sample increased the generalizability of the results, the laboratory context was somewhat far away from a real life experience of yogurt consumption. Several respondents mentioned that they would have reacted more strongly in a real consumption context (i.e., satisfaction scores would have been more positive [negative] in the positive [negative] surprise

condition). Study 4 will aim at improving this aspect. Internal validity of the results, on the other hand, was ensured through the design and matching of the participants.

STUDY 4

Since studies 1 and 2–reporting a direct and indirect influence of surprise on satisfaction–included a wide variety of very different products/services and surprising features, and the results for study 3–using one single product/service and surprising feature–were slightly different, the purpose of study 4 was to replicate study 3 with another product/service and surprising feature. Another purpose of this last study was to use a more realistic setting than in study 3.

Method

As for study 3, the experimental design was a "after only with control group" design including two experimental groups (EG^+ & EG^-) and two control groups (CG^+ & CG^-). The product/service chosen to be bought and consumed during the experiment was a subscription for virtual visits of museums. Respondents–recruited among a non-student population (age between 25 and 60)–were matched according to gender, age, involvement with virtual museums (Zaichowski (1985) scale) and the liking for museums, and then randomly assigned to one of the four conditions. In EG^+, surprise was elicited by the apparition on the computer screen and the printing–after the first virtual visit and the payment of the subscription–of a coupon for a free book. The respondents of EG^+ did not know that they would receive such a coupon. This was supposed to be a pleasant surprise. In EG^-, surprise was elicited by the apparition on the screen–after the first visit and the payment of the subscription–of a message mentioning that they would actually not receive the free book due to out of stock problems. Respondents of EG^- had been told from the start that the subscription also included a free book; this was thus supposed to be an unpleasant surprise. In CG^+, the coupon for the free book was presented as part of the subscription from the start and appeared on the screen and was printed–as announced–after the payment of the subscription. In CG^-, there was no mentioning of a coupon for a free book and it

never appeared on the screen (nor was it printed). Thus, nothing was surprising for the respondents of the two CG. Pre-tests were carried out on 20 respondents from the same non-student population prior to the experiment in order to check the manipulation and the material.

When each respondent arrived in the virtual visit center (in fact the laboratory), the experimenter asked him/her to fill in a mood form and explained the full task[11] while making a demonstration on how to use the program of virtual visit (in order to avoid any surprise related to other elements than the manipulation). Then the respondent was left alone while visiting the virtual museums and answering the final questionnaire (i.e., manipulation check; emotion and satisfaction scales; disconfirmation; and, appreciation of receiving a book). The whole task lasted 75 minutes, including a debriefing.

Exactly the same measures for emotions, disconfirmation and satisfaction as in study 3 were used. A three items (10 points scales) measure of the appreciation/liking of receiving a book and control measures of mood (Peterson and Sauber (1983) scale), extraversion/introversion (Saucier (1994) scale), and expertise with museums (Ohania (1990) scale) were also collected (note that the last three measures had been recorded prior to the day of the experiment).

Results

One hundred and twenty-eight valid questionnaires were collected (32 for each condition). Thirteen questionnaires had to be discarded for the same kind of reasons than for study 3. None of the 64 respondents from the two CG declared having experienced surprise whereas all 64 respondents in the EG were surprised (this surprise was pleasant for EG^+ and unpleasant for EG^-: $mean_{EG+}$ 8.2/10 and $mean_{EG-}$ 3.0/10 on the scale anchored by "very negative (1)" to "very positive (10)"). Highly significant differences were also found between the two EG and the two CG for the non-verbal measures of surprise. Corroborating the work by Reisenzein (2000) and Vanhamme (2000), surprise was–as in study 3–most often indicated by raised eyebrows (44% of the respondents in EG^+ and 28% of the respondents in EG^-) and opened eyes (EG^+: 47%; EG^-: 16%); opened mouth happened the least

often (EG$^+$: 0%; EG$^-$: 12%). Focus of the attention/exploratory behavior was observed for 47% of the respondents in EG$^+$ and for 12% of the respondents in EG$^-$. Finally, a few respondents had a spontaneous surprise exclamation (9% in EG$^+$ and 12% in EG$^-$).

Scales were assessed for uni-dimensionality (using PCA). They were then computed for each group and the distributions were checked for normality. The normality assumption did not hold 1) for verbal surprise for the two CG and 2) for all non-verbal measures (except for OBSS for GE$^+$). Parametrical statistics are thus not appropriate for those variables. Reliability analysis was also performed on all multi-items verbal scales (all Cronbach *alpha* were above .7 except for disconfirmation which was .67). All inter-judges agreements for facial coding and other observed measures were very high (Rust and Cooil (1994) PRL index >.9). As previous studies (e.g., Reisenzein 2000, Vanhamme 2000), non-significant or weak correlations were found between verbal and non-verbal measures of emotions (see table 4). Finally, as expected no difference was found between the two EG and the two CG with respect to the control measures and these measures did not influence the pattern of results (they will thus not be mentioned further).

As expected (hypothesis 2a) all scores of satisfaction–except SATEXP scores–are higher for GE$^+$ than for the two CG, although not always significantly higher for α = .1. The appreciation/liking of receiving the book is also significantly higher for EG$^+$ than CG$^+$ (table 3, lower part).

However, contrary to hypothesis 2b, satisfaction scores were not lower in EG$^-$ than in the CG. The results further show that the scores for the liking of receiving the book seem to be lower in EG$^-$ than in the CG but this difference is not significant.

As for study 3, hypotheses 1a, 1b and 3 were tested by means of correlational analysis between single indicators and the PLS approach for modeling structural equations, mode A of estimation (Wold 1980). As both analyses lead to the same general conclusions, only the results for the PLS approach will be described here (see table 4 for correlations table).

The model including all possible indicators for each of the latent variables in EG$^+$ (i.e., indicators of joy, surprise, disconfirmation and satisfaction) had a good predictive validity (i.e.,

significant Q² and R² for satisfaction) and a good convergent validity for joy (.67>.5) and satisfaction (.91>.5). For surprise, convergent validity was not good enough (.43 < .5). Except for surprise, the model had also good discriminant validity (.51 < .67). The model has thus been re-estimated using only the indicators of surprise that were sufficiently correlated with the estimates of the latent variable. Results are provided in figure 2 (section C). This model is very close to the first model and respects all criteria of predictive (significant Q² and R²), convergent (.59 < .5, .66 < .5 and .91 < .5) and discriminant (.51 < .59) validity. The analysis of this model shows that hypothesis 1a is only partially supported by the data in that the direct influence of surprise on satisfaction seems to be fully mediated by joy. The path coefficient between the latent variables of surprise and satisfaction is not significant although the correlation between these latent variables was significant (for α = .07); analysis show that this is due to joy. Furthermore, as expected (hypothesis 3), the (indirect) influence of surprise on satisfaction remains significant despite the presence of disconfirmation in the model.

The model including all possible indicators for each of the latent variables in the EG⁻ (i.e., indicators of negative emotions, surprise, disconfirmation and satisfaction) also suffered from a problem of convergent validity for surprise. In this model, the only significant correlations between latent variables were between surprise and negative emotions and between disconfirmation and satisfaction. The path coefficients between these variables were also significant. No other significant link appeared in this model. The results are the same for the new model–including only the indicators of surprise that were sufficiently correlated with the estimates of the latent variable–which satisfies all criteria of predictive, convergent and discriminant validity. Thus, the data do not corroborate hypotheses 1b and 3. It should be noted that the PLS model and the correlations pattern between verbal and nonverbal indicators of emotions indicated–as in study 3–the probable usage of display rules (correlations between verbal and non-verbal measures of emotions were negative in the negative surprise condition whereas they were positive for the positive surprise condition, see table 4).

Conclusion and Discussion

On the one hand, the results show that pleasantly surprised consumers reported higher satisfaction scores than non-surprised consumers, even though the results were not highly significant (which could partially be explained by the small sample sizes, i.e., 32). Consumers also evaluated more positively the manipulated element (i.e., the coupon for the book) in the positive surprise condition than in the non-surprise condition. In this study, surprise seemed to influence satisfaction only indirectly, through its amplification of the emotion of joy–the direct influence was totally mediated by joy. Finally, disconfirmation did not capture the influence of surprise on satisfaction.

For negatively surprised consumers, on the other hand, satisfaction scores were not lower than for non-surprised consumers and no link appeared between surprise and satisfaction within the negative surprise condition. With respect to the free book, the scores of liking seemed lower in the negative surprise condition than in the control group but this difference was not significant and could thus be due to chance.

The fact that surprise did apparently not influence satisfaction in the negative condition whereas it had a significant influence in the positive surprise condition could be due to the nature of the objet that was used to manipulate surprise. The book is not necessary for the delivery of the core service, i.e., the virtual visit. And the performance of this core service was very good. Therefore, the negative surprise elicited vis-à-vis this non-necessary object might have been considered of little relevance for the evaluation of the whole consumption experience. This explanation is consistent with the *How-do-I-feel-about-it?* heuristic.

With respect to the limitations of the study, the setting was much more natural than for study 3. During the debriefing, none of the respondents (all non-students as for study 3) mentioned that they would have reacted differently in a real context except for six respondents in the negative surprise condition (as for study 3, those respondents said they would have had lower satisfaction scores in a

real context). Compared to study 3, external validity was thus enhanced. Internal validity of the results was also ensured through the design and matching of the respondents as for study 3.

GENERAL DISCUSSION: CONTRIBUTIONS, LIMITATIONS AND FUTURE PATHS OF RESEARCH

A better knowledge of the influence of surprise is not only necessary for researchers but also for managers who want to use (or already use) surprise as a marketing tools (e.g., some companies send a small gift on their client's birthday, or give them once in a while tickets to attend a match of their favorite sport). Some studies show that companies using surprise following their gut feeling–without knowing anything about this emotion and its effects–sometimes face disastrous consequences (e.g., Vanhamme and Lindgreen 2001).

A major contribution of this research thus concerns the place of surprise in models of satisfaction formation. All studies carried out in this research concluded that surprise is worth studying on its own and that its influence on satisfaction is not fully captured by disconfirmation. They also re-emphasize the importance of enriching cognitive models with emotional variables. Both field studies revealed a significant direct and indirect relationship between surprise and satisfaction (positive / negative emotions did not fully mediate the impact of surprise on satisfaction). These field studies thus led to convergent results despite their use of different methods of data collection, which ensures some generalization of the results. For the two laboratory experiments, the pattern of results is, however, not as clear-cut. For the positive surprise condition of the first experiment, surprise only had a significant positive direct impact on satisfaction, whereas this impact was totally mediated by joy for the second experiment. Moreover, positively surprised customers of the first experiment were not more satisfied than the non-surprised customers, whereas the positively surprised customers of the second experiment were more satisfied than the non-surprised customers. Furthermore, for both experiments, scores for the evaluation of the surprising element were higher in the positive surprise

condition that in the non-surprise condition. For the negative surprise condition, negatively surprised respondents of the first experiment had lower satisfaction scores and lower evaluations of the surprising element than the non-surprised consumers. Furthermore, surprise had both a significant negative direct and indirect impact on satisfaction. However, no impact at all on satisfaction was found in the second experiment.

Satisfaction studies have shown that models of satisfaction formation may vary according to the type of product/service (e.g., Churchill and Surprenant 1982). On the one hand, the field studies– that revealed both a direct and indirect relationship (correlation) between surprise and satisfaction– were based on a large set of different product/service experiences. Each experiment, on the other hand, was only based on one single product/service. This could be a first explanation for the differences. [16] The experiments should thus be replicated with other products/services. How the type of relationship (direct and/or indirect) varies according to types of product/service and/or surprising feature could then be checked.

Moreover, the different results that appeared between studies 3 and 4 for the positive surprise condition could be explained by a difference in the 'base level' of satisfaction (i.e., satisfaction level without effect of surprise): respondents from the non-surprise conditions in the first experiment had much lower scores of satisfaction in average than those of the second experiment. There might thus exist a critical 'base level' of satisfaction that would need to be reached–prior to any manipulation of

---

[16] Note also that for studies 1 and 2, the type of judgment (on-line, memory-based or inference-based) could not be controlled whereas satisfaction judgements were most probably formed on-line for studies 3 and 4 (see note 11). Hypotheses 1 to 3 do not depend on the type of process. However, recalled material used to form a memory-based judgement might suffer from some bias that does not exist for on-line judgements. On-line judgements should be the least biased as the only recalled material in t+1 is the satisfaction judgement itself whereas all the components necessary to form the satisfaction judgement (e.g. recalled emotions, inferences, etc.) have to be recalled for memory- or inference-based judgments (Vanhamme and Snelders 2001). These additional biases might also partially be responsible for the differences obtained between the field studies and the experimental studies. This aspect should thus be further investigated.

surprise–in order to secure a significant difference in satisfaction scores due to surprise (in experiment one, this critical base level was not reached). The paths of influences (direct or indirect) could also be influenced by this critical base level.

With regard to the negative surprise conditions of the two experiments, the nature of the element that was used for surprising the respondent might explain the differences in the results. The spoon in the negative surprise condition was considered to be a necessary element for eating the yogurt ($\approx$ core service), whereas the free book was totally unnecessary to have in being able to visit the museum ($\approx$ core service). A negative surprise with a necessary element for the production of the core service may thus have been considered as more relevant for the evaluation of the satisfaction with the whole consumption experience than a non-necessary element. This would explain the absence of significant effect for the negative surprise condition in the second experiment.

Replication of these experiments should thus also try to investigate the moderating impact of 1) the satisfaction base level, and 2) the nature of the element used for eliciting surprise. Other moderators might also be considered, such as, for example, 'surprise aversion'. Some people might not like to be surprised. A supposedly pleasant surprise may therefore lead to a negative impact on satisfaction as far as those consumers are concerned. The concept of *surprise-aversion* has not yet been investigated in the literature. However, 'surprise averse' consumers are likely to be characterized by a willingness to plan, control, and master everything, and by a low optimal stimulation level. This concept remains to be investigated further.

A limitation of the two experimental studies is that some respondents reported they would have reacted differently–more strongly–in a real consumption context. This was the case for the negative surprise conditions of both experiments and for the positive surprise condition of the first experiment. The results for the positive surprise condition of the second experiment–i.e., the only experimental condition for which no respondent reported s/he would have reacted differently in reality–are the closest to the results reported for the field studies. The context might thus partly

explain the different results obtained for the field studies and laboratory studies. Replications of these experimental studies should thus strive to use as natural a setting as possible.

It is, however, worth noting that construct validity of the satisfaction measures should be higher for the experimental studies than for studies 1 and 2 (since these measures were recorded right after the consumption/purchase) and that these studies improved upon most other studies dealing with emotions in that it used truly different measures for emotions (and surprise in particular).

It should be emphasized that the purpose of the present research was to investigate the impact of surprise on a short-term satisfaction, i.e., on the transaction-specific satisfaction. Some other types of satisfaction exist, such as, cumulative satisfaction, which adopt a longer-term perspective. It would thus be worthwhile investigating the kind of influence surprising consumption / purchase experiences have in the long run. Moreover, this study only considered post-purchase marketing variables. Surprise could, however, also have an impact on pre-purchase variables. Managers will then have to decide whether they want to surprise consumers on the point of sales (for example) in order to trigger a purchase (e.g., by posting a message that the product contains a gift) or after they have purchased the product/service in order to increase satisfaction and retention. Both are usually exclusive: in the latter case, managers will not be allowed to reveal that there is a gift in the product in order to trigger the purchase.

As a result of the present research, surprise indeed appears to be another tool that can be used by marketers to increase customer satisfaction. From a managerial point of view, it is worth thinking thoroughly about how to elicit pleasant surprise because such surprised consumers will not only be more satisfied but will probably also repurchase from the company and advertise their experience–through word-of-mouth. It is also important–for companies using surprise as a marketing tool–to communicate properly about their 'surprise' strategy. If they do not communicate well and let the consumers update their schema, they are not likely to surprise them once more with the same trick, and will probably increase their expectations for their next purchase. However, if they can entertain suspense (Pine and Gilmore 1999), they will be able to keep surprising their customers and avoid that

danger. For example, some companies send reduction coupons to their customers following a systematic pattern (e.g., 20% off in March, 30% off in May, and 40% off in June). After having experienced it a few times, customers very often update their schema and wait for the highest reduction coupon before purchasing. As a result, not only are they not surprised anymore when they receive the coupon but also they stop purchasing without receiving a reduction coupon (this is likely to lead to a reduction in profit for the company). The same kind of situation happens for frequent flyer programs. The first time a gold member is up-graded from economy-class to business-class s/he is likely to be surprised but then s/he might consider it as one of his/her privileges (schema up-date) and believe the company should upgrade her/him for all her/his trips. Such schema up-date could be avoided if the company communicated properly about the 'surprise feature'. For example, by letting the customer believe it is due to chance (e.g., lottery) or linked to an occasional event (e.g., the 75th birthday of the company).

Surprising customers is not free of costs; companies need to invest time in understanding their customers and in creating surprising aspects in–or around–the product/service delivered. The element used to surprise does not, however, need to be expensive, little and cheap things can work (e.g., a birthday card, a free ticket for a show). It is, however, worth noting that it can be difficult–for mass product consumption–to ensure that all consumers will be surprised by the same aspect (because surprise depends on the consumers' individual schemas). In one-to-one relationships with the customers, companies have a better knowledge of their customers' schemas and are, therefore, more likely to succeed in surprising them. A careful analysis of costs and benefits is thus recommended. Furthemore, when using surprise as a marketing tool, companies should adhere to clear ethical guidelines: companies blindly guided by their profitability might misuse surprise.

To conclude, it should be pointed out that–due to the 'rich' properties of surprise (see figure 1)–the conceptual framework of surprise and the investigation of its influence, developed in the present research, would be worth extending to marketing variables other than satisfaction, such as word-of-mouth, attitude toward the ad, ad effectiveness, memorization of ads and brand names.

APPENDIX A. MAIN SCALES USED IN THE FOUR STUDIES

Izard's DES (1977) items for surprise, joy and negative emotions (anger, sadness, fear, disgust and disdain) (5 points scale items anchored from "not at all" (1) to "a lot (5)"):

Surprise - 3 items: surprised, amazed, astonished

Enjoyment - 3 items: joyful, delighted, happy

Distress - 3 items: sad, downhearted, discouraged

Anger - 3 items: angry, mad, enraged

Fear - 3 items: afraid, scared, fearful

Disgust - 3 items: disgusted, feeling of distaste, feeling of revulsion

Contempt/disdain - 3 items: disdainful, contemptuous, scornful

Items for disconfirmation (Oliver 1980, 1997)

* 1) Overall this X was: *Respondents had to circle either 1[much worse than expected], 2 [worse than expected], 3 [as expected], 4 [better than expected] or 5 [much better than expected]*

2) Considering only the benefits/positive aspects received from this X, were they: *Respondents had to circle either 1[much worse than expected], 2 [worse than expected], 3 [as expected], 4 [better than expected] or 5 [much better than expected])*

3) Considering only the problems /negative aspects encountered with this X, were they: *Respondents had to circle either 1[much worse than expected], 2 [worse than expected], 3 [as expected], 4 [better than expected] or 5 [much better than expected]*

[*] only item used in studies 1 and 2; the third item had to be discarded for studies 3 and 4 since it strongly reduced the reliability.

DT scale

Delighted	Pleased	Mostly satisfied	Mixed (about equally satisfied and dissatisfied)	Mostly dissatisfied	Unhappy	Terrible
(7)	(6)	(5)	(4)	(3)	(2)	(1)
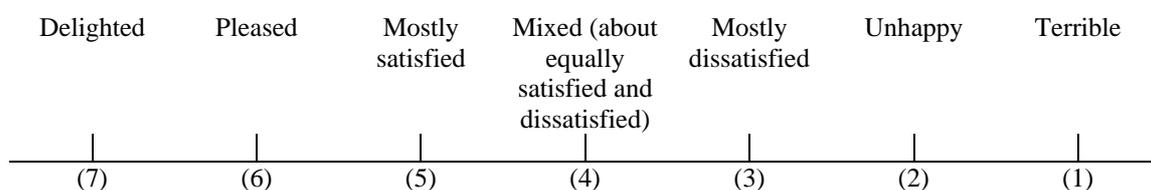
Items for SAT (Oliver 1980, 1997)–10 points items anchored from 1 (totally disagree) to 10 totally agree

1) This is one of the best X I could have bought

2) This X is exactly what I need

3) This X was not as good as I thought it would

* 4) I am satisfied with X

5) Sometimes I have mixed feelings about keeping X [1]

* 6) My choice to buy X was a wise one [1]

* 7) If I could do it over again, I'd buy a different X [1]

8) I have truly enjoyed X

* 9) I feel bad about my decision to buy this X [1]

* 10) I am not happy that I bought ['ate', for study 3] X

11) Owning ['consuming' for study 3] X has been a good experience

* 12) I am sure it was the right thing to buy this X [1]

(*) Items from the 6-items Oliver (1980)'s scale (6-items scale used in the field studies).

(1) These items were left out in study 3 as they did not fit the context of the experiment.

Items for SATEXP (Aurier & Evrard (1998))–studies 3 and 4

1) How satisfied are you with X? *Respondents had to circle a number between 1(X did not please me at all) and 10 (X pleased me very much).*

2) Compared to other X you bought before, this X was: *Respondents had to circle a number between 1 [the worst] and 5 [the best])–this item was left out in study 4 as it did not fit the context of the experiment.*

3) To what extent does X resemble your ideal X? *Respondents had to circle either 1[to a very large extent], 2 [to a large extent], 3 [to an average extent], 4 [to a lesser extent] or 5 [to a very small extent].*

4) Would you recommend X to your friends? *Respondents had to circle either 1[definitely yes], 2 [probably yes], 3 [perhaps], 4 [probably not] or 5 [definitely not].*

REFERENCES

Anderson, Eugene W. and Claes Fornell (1994), "Customer Satisfaction Prospectus", in *Service Quality: New Directions in Theory and Practice*, eds. Rust Roland T. and Richard L. Oliver, London: Sage, 241-268.

Aurier, Philippe and Yves Evrard (1998), "Elaboration et Validation d'une Echelle de Mesure de la Satisfaction des Consommateurs", in *Actes de la Conférence de l'Association Française du Marketing*, *14, 1,* Saporta, Bernard and Jean-françois Trinquecoste, Bordeaux: IAE, 51-71.

Bitner, Mary Jo (1990), "Evaluating Service Encounters: the Effects of Physical Surroundings and Employee Responses", *Journal of Marketing*, 54 (2), 69-82.

Cantor, Joanne R., Dolf Zillmann, and Jennings Bryant (1974), "Enhancement of Experienced Sexual Arousal in Response to Erotic Stimuli Through Misattribution of Unrelated Residual excitation", *Journal of Personality and Social Psychology*, 32 (1), 69-75.

Charlesworth, William R. (1969), "The Role of Surprise in Cognitive Development", in *Studies in Cognitive Development, Essays in Honor of Jean Piaget,* eds. Elkind, David and Flavel, J.H., London: Oxford University Press, 257-314.

Choi, Kyungmee and John Marden, J (1998), "A Multivariate Version of Kendall's tau"*, Journal of Nonparametric Statistics*, 9 (3), 261-293.

Churchill, Gilbert A. (1999), *Marketing Research: Methodological Foundations*, 7th edition, Toronto: The Dryden Press.

Churchill, Gilbert A. and Carol Surprenant (1982), "An Investigation into the Determinants of Customer Satisfaction", *Journal of Marketing Research 19* (November), 491-504.

Cohen, Joel B and Charles S. Areni (1991), "Affect and Consumer Behavior", in *Handbook of Consumer Theory and Research,* eds. Robertson, Thomas S. and Harold H. Kassarjian, Englewood Cliffs, NJ: Prentice Hall, 188-240.

Desai, Mahesh M. (1939), "Surprise: A Historical and Experimental Study", *British Journal of Psychology, Monogr. Suppl.*, 22, 1-124.

Derbaix, Christian and Ingrid Poncin (1999), "Process Tracing of Affective Reactions elicited by Commercials: Assessing the Explanatory Power of the Outputs of the Feelings Monitor and of the Coding of Facial Expressions", in *European Advances in Consumer Research,* Volume 4, eds. Bernard Dubois, Lowry, T. and Shrum, L., 166-173.

Derbaix, Christian and Michel Tuan Pham (1989), "Pour un Développement des Mesures de l'Affectif en Marketing", *Recherche et Applications en Marketing*, 4 (4), 71-87.

Dijkstra, Taeke Klaes (1981), *Latent Variables in Linear Stochastic Models: Reflections on "Maximum Likelihood" and "Partial Least Squares" Methods*, Ph.D. Dissertation, Rijksuniversiteit te Groningen.

Ekman, Paul and Wallace V. Friesen (1975), *Unmasking the Face,* N-J: Prentice Hall.

Fiske, Susan T. (1980), "Attention and Weight in Person Perception: the Impact of Negative Extreme Behavior", *Journal of Personality and Social Psychology*, 38 (6), 889-906.

Flanagan, John C. (1954), "The Critical Incident Technique", *Psychological-Bulletin,* 51, 327-358.

Fornell, Claes and Fred L. Bookstein (1982), "Two Structural Equation Models: LISREL and PLS applied to Consumer Exit-voice Theory", *Journal of Marketing Research*, 9 (November), 440-452.

Gorn, Gerald, Michel Tuan Pham, and Leo Y. Sin (2001), "When Arousal influences Ad Evaluation and Valence Does Not (and Vice Versa)", *Journal of Consumer Psychology*, 11 (1), 43-55.

Izard, Caroll E. (1977), *Human Emotions*, New-York: Plenum Press.

Jain, Kapil and Narasimhan Srinivasan (1990), "An Empirical Assessment of Multiple Operationalizations of Involvement", *Advances in Consumer Research*, 17, 594-602.

La Barbera, Priscillia A. and David Mazursky (1983), "A Longitudinal Assessment of Consumer Satisfaction/Dissatisfaction: The Dynamic Aspect of Cognitive Standardization", *Journal of Marketing Research,* 20 (4), 394-404.

Mano, Haim and Richard L. Oliver (1993), "Assessing the Dimensionality of Consumption Experience: Evaluation, Feeling, and Satisfaction", *Journal of Consumer Research*, 20 (3), 451-466.

Mellers, Barbara, Alan Schwartz, and Ilana Ritov (1999), "Emotion-Based Choice", *Journal of Experimental Psychology: General,* 128, 3, 332-345.

Meyer, Wulf-Uwe, Rainer Reisenzein, and Achim Schützwohl (1997), "Toward a Process Analysis of Emotions: The Case of Surprise, *Motivation and Emotion,* 21 (3), 251-274.

Ohania (1990), "Scale of Expertise", in *Handbook of Marketing Scales : Multi-item Measures for Marketing and Consumer Behavior Research*, eds. Bearden, William O. and Richard G. Netemeyer (1999), Thousand Oaks: Sage.

Oliver, Richard L. (1980), "A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions", *Journal of Marketing Research*, 17 (4), 460-469.

_____(1981), "Measurement and Evaluation of Satisfaction Processes in Retail Settings", *Journal of Retailing,* 57 (3), 49-68.

_____ (1993), "Cognitive, Affective and Attributes Bases of the Satisfaction Response" *Journal of Consumer Research*, 20 (3), 418-430.

_____ (1994), "Conceptual Issues in the Structural Analysis of Consumption Emotion, Satisfaction and Quality: Evidence in a Service Setting", in *Advances in Consumer Research*, 21, eds. Allen, Chris T. and Deborah Roedder John, Provo, UT: Association for Consumer Research, 16-22.

_____(1997), *Satisfaction: on a Behavioral Perspective on the Consumer*, New-York: McGraw-Hill.

_____and Wayne S. de Sarbo (1988), "Response Determinants in Satisfaction Judgements", *Journal of Consumer Research* ,14 (March), 495-507.

_____, Roland T. Rust, and Sajeev Varki (1997), "Customer Delight: Foundations, Findings, and Managerial Insight. *Journal of Retailing,* 73 (3), 311-336.

_____ and Robert A. Westbrook (1993), "Profiles of Consumer Emotions and Satisfaction in Ownership and Usage", *Journal of Consumer Satisfaction*, *Dissatisfaction, and Complaining Behavior,* 6, 12-27.

Peterson, Robert A. and Matthew Sauber (1983), "Mood Short Form: MSF", in *Handbook of Marketing Scales : Multi-item Measures for Marketing and Consumer Behavior Research*, eds. Bearden, William O. and Richard G. Netemeyer (1999), Thousand Oaks: Sage.

Pham, Michel Tuan (1998), "Representativeness, Relevance, and the Use of Feelings in Decision Making", *Journal of Consumer Research*, 25 (2), 144-159.

_____, Joel B. Cohen, John. W. Pracejus, and G. David Hughes (2001), "Affect Monitoring and the Primacy of Feelings in Judgment", *Journal of Consumer Research*, 28 (2), 167-188.

Phillips, Diane M. and Baumgartner Hans (2002), "The Role of Consumption Emotions in Satisfaction Response, *Journal of Consumer Psychology*, 12 (3), 243-252.

Pine, Joseph B. and James H. Gilmore (1999). *The Experience Economy: Work is Theatre and Every Business a Stage,* Boston: Harvard business school press.

Plutchik, Robert (1980), *Emotion: a Psychoevolutionary Synthesis*, New-York: Harper and Row.

Price, Linda L., Eric J. Arnould and Deibler, S. L. (1995), "Consumers'Emotional Responses to Service Encounters", *International Journal of Service Industry Management,* 6 (3), 34-63.

Rao, Sudhakar H. and Gerald L. Sievers (1995), "A Robust Partial Correlation Measure", *Journal of Nonparametric Statistics*, 5 (1), 1-20.

Reisenzein, Rainer (1983), "The Schachter Theory of Emotion ", *Psychological Bulletin*, 94 (2), 239-264.

_____ (2000), "Exploring the Strength of Association between the Components of Emotion Syndrome: the Case of Surprise", *Cognition and Emotion,* 14 (1), 1-38.

_____ and Sandra Bördgen (1998), *Evidence for Strong Dissociation Between Emotion and Facial Display: The Case of Surprise*, unpublished manuscript, Faculty of Psychology, University of Bielenfeld, Bielenfeld, Germany.

Richins, Marcha L. (1997), "Measuring Emotions in the Consumption Experience", *Journal of Consumer Research*, 24 (2), 127-146.

Rumelhart, David E. (1984). "Schemata and the Cognitive System", in *Hanbook of Social Cognition 1,* eds. Wyer, R.S. Jr. and Srull, T.K., Hillsdale, NJ: Lawrence Erlbaum Associates Inc, 161-188.

Rust, Roland T. and Bruce Cooil (1994), "Reliability Measures for Qualitative Data: Theory and Implications", *Journal of Marketing Research*, 31 (February), 1-14.

_____and Oliver, R.L. (2000), "Should We Delight the Customer ?", *Journal of The Academy of Marketing Science,* 28 (1), 86-94.

_____, Zahorik, A. J., and Keiningham, T. L. (1996), *Service marketing,* New York, Harper Collins.

Saucier, Gerard (1994), "Mini-markers: a Brief Version of Goldberg's Unipolar Big Five Markers", *Journal of Personality Assessment*, 63 (3), 506-516.

Scherer, Klaus R. (1984), "On the Nature and Function of Emotion: a Component Process Approach", in *Approaches to emotion*, eds. Scherer, K.R. and Ekman, P., Hillsdale, NJ: Lawrance Erlbaum Associates Inc., 293-317.

Schneider, Benjamin and David E. Bowen (1999), "Understanding Customer Delight and Outrage", *Sloan Management Review*, 41(1), 35-45.

Schützwohl Achim (1998), "Surprise and Schema Strength", *Journal of Experimental Psychology, Learning, Memory and Cognition,* 24(5), 1182-1199.

Schwarz Norbert (1990), "Feelings as Information: Informational and Motivational Functions of Affective States"*, Handbook of Motivation and Cognition: Foundations of Social Behaviour*, 2, eds. Higgins, E. Tory and Sorentino, R., New-York: Guilford, 527-561.

Schwarz, Norbert and Gerald L. Clore (1996), "Feelings and Phenomenal Experiences", in *Social Psychology: Handbook of Basic Principles*, eds. Higgins, E. Tory and Kruglanski, Arie W., New York: Guilford press, 433-465.

Siegel, Sidney and N. John Castellan, Jr. (1988), *Nonparametric Statistics for the Behavioral Sciences*, New-York: MacGraw Hill.

Stone, Arthur A., Ronald C. Kessler, and Jennifer A. Haythornthwaite (1991), "Measuring Daily Events and Experiences: Decisions for the Researcher", *Journal of Personality*, 59(3), 575-607.

Tenenhaus, Michel (1998), *La Régression PLS: Théorie et Pratique*, Paris: Editions Technip.

Valette-Florence, Pierre (1988), Analyse Structurelle Comparative des Composantes des Systèmes de Valeurs selon Kahle et Rokeach, *Recherche et Applications en Marketing*, 3 (1), 15-34.

Vanhamme, Joëlle (2000), "The Link Between Surprise and Satisfaction: an Exploratory Research on How Best to Measure Surprise", *Journal of Marketing Management*, 16 (6), 565-582.

_____ and Adam Lindgreen (2001), "Gotcha! Findings from an exploratory investigation on the dangers of using deceptive practices in the mail order business", *Psychology & Marketing*, 18 (7), 785-810.

_____ and Dirk Snelders (2001). The role of Surprise in Satisfaction Judgments? *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior,* 14, 27-45.

Westbrook, Robert A. (1980), "A Rating Scale for Measuring Product/Service Satisfaction, *Journal of Marketing*, 44 (4), 68-72.

_____ (1987), "Product/consumption-based Affective Responses and Postpurchase Process", *Journal of Marketing Research*, 24, 258-270.

_____ and Richard L. Oliver (1991), "The Dimensionality of Consumption Emotion Patterns and Consumer Satisfaction, *Journal of Consumer Research*, 18 (1), 84-91.

Wheeler, Ladd and Harry T. Reis (1991), Self-recording of Everyday Life Events: Origins, Types and Uses", *Journal of Personality*, 59 (3), 339-354.

Wold, Herman (1980), "Model Construction and Evaluation When Theoretical Knowledge is Scarce: Theory and Applications of Partial Least Squares", in *Evaluation of econometric models*, eds. Kmenta, Jan. and James B. Ramsey, New York: Academic Press, 47-74.

Zaichkowsky, Judith L. (1985), Measuring the Involvement Construct, *Journal of Consumer Research,* 12 (December), 341-352.

Zillmann, Dolf (1983), "Transfer of Excitation in Emotion Behavior", in *Social Psychophysiology*, eds. Cacioppo, John T. and Richard A. Petty, New-York: Guilford, 215-240.

TABLE 1

KENDALL CORRELATIONS AND PARTIAL CORRELATIONS

STUDY 1

Section A. *Kendall and Pearson correlations–negatively-valenced experiences*

		Anger	Fear	Disgust	Surprise
Kendall *T*	Surprise	.584** (n = 50)	.321** (n = 50)	.494** (n = 50)	
	DT	-.515** (n = 49)	-.258* (n = 49)	-.452** (n = 49)	-.501** (n = 49)
	SAT	-.314** (n = 46)	-.210* (n = 46)	-.228* (n = 46)	-.406** (n = 46)

Section B. *Kendall and Pearson partial correlations–negatively-valenced experiences*

	NE$_A$[a]	NE$_B$[b]	Anger	Fear	Disgust
Accounted for[d]:					
T DT-surprise	-.44**	-.31**	-.29**	-.46**	-.36**
T SAT-surprise	-.37**	-.32**	-.29**	-.37**	-.35**

STUDY 2

Section C. *Correlations - negatively surprising experiences (N = 50)*

		Anger	Sadness	Fear	Disgust	Disdain	Surprise
Kendall *T*	Surprise	*n.s.*	.246*	.288*	.272*	.192*	
	DT	-.383**	*n.s.*	-.218*	-.329**	-.210*	*n.s.*
	SAT	*n.s.*	*n.s.*	*n.s.*	*n.s.*	-.237*	*n.s.*

Section D. *Correlations - negative sample (N = 100)*

		Anger	Sadness	Fear	Disgust	Disdain	Surprise
Kendall *T*	Surprise	.615**	.616**	.548**	.645**	.604**	
	DT	-.749**	-.620**	-.507**	-.684**	-.577**	-.607**
	SAT	-.572**	-.530**	-.419**	-.550**	-.520**	-.546**

Section E. *Partial correlations - negative sample (N = 100)*

	NE$_C$[c]	Anger	Sadness	Fear	Disgust	Disdain
Accounted for[d]:						
T DT-surprise	-0,28**	-0,28**	-0,36**	-0,46**	-0,30**	-0,40**
T SAT-surprise	-0,28**	-0,30**	-0,33**	-0,42**	-0,30**	-0,34**

Legend:** $p_{1-tail} < .01$; * $p_{1-tail} < .05$

Notes:

[a] Mean for all items of negative emotions.
[b] Mean of items of negative emotions that have high loading on the first factor (PCA - one factor solution).
[c] Mean of all items of negative emotions (they all loaded very highly on the first factor of PCA - one factor solution).
[d] The partial Kendall correlation only allows to control for one variable at a time.

TABLE 2

PEARSON (LOWER TRIANGLE) AND KENDALL (UPPER TRIANGLE) CORRELATIONS; TEST OF H1A,B AND H3 FOR VERBAL SURPRISE - 1 TAIL TESTS - STUDY 3

GE^+ (n = 31)

	DT	SAT	SATEXP	VSUR	OBSS	FS	BS	EYE	EYEB	MOUTH	EXCL	STOP	EXPL	VJ	FJ
DT		.547***	.510***	.342*	.226•	.382**	n.s	n.s.	.370**	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
SAT	.685***		.738***	.293*	.214*	.316*	n.s	.277*	.417**	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
SATEXP	.660***	.900***		.290*	n.s	.295*	n.s	.260*	.367**	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
VSUR	.383*	.379*	.333*		n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	.223•	n.s.
OBSS	.324*	.358*	.307*	n.s.		.641***	.436**	.423**	.569***	.332*	.409**	.248•	.401**	.424***	.325*
FS	.343*	.371*	.345*	n.s.	.726***		n.s.	.608***	.870***	.389**	.479**	n.s.	.234°	.328*	.358*
BS	n.s.	n.s.	n.s.	n.s.	.586***	n.s.		n.s.	n.s.	n.s.	n.s.	.616***	.710***	.279*	n.s.
EYE	n.s.	.312*	.287•	n.s.	.496**	.833***	n.s.		.554***	.329*	.552***	n.s.	n.s;	n.s.	n.s.
EYEB	.377*	.509**	.475**	n.s.	.679***	.882***	n.s.	.683***		n.s.	.231°	n.s.	n.s.	.234•	n.s.
MOUTH	n.s.	n.s.	n.s.	n.s.	.398*	.583***	n.s.	.460**	n.s.		.795***	n.s.	n.s.	.265*	.740***
EXCL	n.s.	n.s.	n.s.	n.s.	.510**	.631***	n.s.	.638***	.304*	.794***		n.s.	.226°	n.s.	.597***
STOP	n.s.	.242°	n.s.	n.s.	.320*	n.s.	.616***	n.s.	n.s.	n.s.	n.s.		.108	n.s.	n.s.
EXPL	n.s.	n.s.	n.s.	n.s.	.552***	n.s.	.710***	n.s.	n.s.	n.s.	n.s.	.302*		n.s.	n.s.
VJ	n.s.	n.s.	n.s.	.279•	.580***	.330*	.384*	n.s.	.262°	.283•	.251°	n.s.	n.s.		.284*
FJ	n.s.	n.s.	n.s.	n.s.	.362*	.375*	n.s.	n.s.	n.s.	.667***	.534***	n.s.	n.s.	0.309*	

Relationship	Mediation tests for verbal surprise (potential moderator is VJ)	full mediation?
		NO
VSUR – DT:	$VJ = f(VSUR) \Rightarrow P_{1\text{-tail}}$ .064 ; $DT = f(VSUR) \Rightarrow P_{1\text{-tail}}$.017; $DT = f(VSUR, VJ) \Rightarrow VSUR$: $P_{1\text{-tail}}$ .025, VJ: $P_{1\text{-tail}}$ n.s.; Pearson $r_{\text{VSUR-DT.VJ}} = .362^*$ (Kendall $T_{\text{VSUR-DT.VJ}} = .324^•$)	NO
VSUR – SAT:	$VJ = f(VSUR) \Rightarrow P_{1\text{-tail}}$ .064 ; $SAT = f(VSUR) \Rightarrow P_{1\text{-tail}}$ .018 ; $SAT = f(VSUR, VJ) \Rightarrow VSUR$: $P_{1\text{-tail}}$ .030, VJ: $P_{1\text{-tail}}$ n.s.; Pearson $r_{\text{VSUR-SAT.VJ}} = .347^*$ (Kendall $T_{\text{VSUR-SAT.VJ}} = .272^•$)	NO
VSUR – SATEXP:	$VJ = f(VSUR) \Rightarrow P_{1\text{-tail}}$ .064 ; $SATEXP = f(VSUR) \Rightarrow P_{1\text{-tail}}$ .034 ; $SATEXP = f(VSUR, VJ) \Rightarrow VSUR$: $P_{1\text{-tail}}$ .035, VJ: $P_{1\text{-tail}}$ n.s.; Pearson $r_{\text{VSUR-SATEXP.VJ}} = .335^*$ (Kendall $T_{\text{VSUR-SATEXP.VJ}} = .288^*$)	NO

GE^- (n = 34)

	DT	SAT	SATEXP	VSUR	OBSS	FS	BS	EYE	EYEB	MOUTH	EXCL	STOP	EXPL
DT		.444**	.526***	-.310*	n.s.	.206•	n.s.	n.s.	n.s.	n.s.	.276*	n.s.	n.s.
SAT	.537**		.656***	-.170°	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	.241*	n.s.	n.s.
SATEXP	.599***	.815***		n.s.	n.s.	.177°	n.s.	n.s.	n.s.	n.s.	.269*	n.s.	n.s.
VSUR	-.309*	n.s.	n.s.		n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
OBSS	n.s.	n.s.	n.s.	n.s.		.699***	.272*	.401**	.439***	.469***	.503***	.288*	.206°
FS	n.s.	n.s.	n.s.	n.s.	.808***		.187°	.638***	.598***	.432***	.421**	.403**	.198°
BS	n.s.	n.s.	n.s.	n.s.	.458**	n.s.		n.s.	n.s.	n.s.	n.s.	n.s.	.602***
EYE	n.s.	n.s.	n.s.	n.s.	.485**	.795***	n.s.		.567***	.252*	n.s.	.470**	n.s.
EYEB	n.s.	n.s.	n.s.	n.s.	.544***	.713***	n.s.	.652***		n.s.	n.s.	.325*	n.s.
MOUTH	n.s.	n.s.	n.s.	n.s.	.534***	.570***	n.s.	.341*	n.s.		.714***	n.s.	n.s.
EXCL	.262•	.289*	.342*	n.s.	.555***	.503***	n.s.	n.s.	n.s.	.726***		n.s.	n.s.
STOP	n.s.	n.s.	n.s.	n.s.	.356*	.475**	n.s.	.504***	.349*	n.s.	n.s.		n.s.
EXPL	n.s.	n.s.	n.s.	n.s.	.312*	.230°	.602***	n.s.	n.s.	n.s.	n.s.	n.s.	

T Kendall $_{\text{VJ-FJ}} = -.270$, $p_{1\text{-tail}} = .037$; Pearson $r_{\text{VJ-FJ}} = -.320$, $p_{1\text{-tail}} = .033$; N = 34.

Pearson r $_{\text{VSUR-VNE}}$. $= 433$, $p_{1\text{-tail}} < .01$; Kendall $T_{\text{VSUR-VNE}}$ 0.383, $p_{1\text{-tail}} < .01$; N = 34

Pearson correlations for non-normal variables have to be considered with caution; Legend: $^{***}p_{1\text{-tail}} < .001$; $^{**}p_{1\text{-tail}} < .01$; $^*p_{1\text{-tail}} < .05$; $^•p_{1\text{-tail}} < .07$; $^°p_{1\text{-tail}} < .01$;

TABLE 3

SCORE COMPARISON BETWEEN EG AND CG

STUDY 3

	Mean EG+	Mean CG+	Mean CG-	Mean EG-	EG+ versus CG+ (*t* test & WSRT)	EG- versus CG+ (*t* test & WSRT)	EG- versus CG- (*t* test & WSRT)
SPOON	7.14	5.98	-	5.33	$t(30) = 2.61^{**}$ $z = -2.39^{**}$	$t(30) = -1.32^{°}$ $z = -1.06^{(p=.14)}$	-
DT	4.58	4.60	4.58	4.13	*n.s.*	$t(30) = -1.97^{*}$ $z = -1.90^{*}$	$t(33) = -2.45^{*}$ $z = -1.90^{*}$
SAT	5.66	6.03	5.20	4.82	*n.s.*	$t(30) = -2.19^{*}$ $z = -2.03^{*}$	*n.s.*([a])
SATEXP	6.13	6.09	5.81	5.26	*n.s.*	$t(30) = -1.93^{*}$ $z = -1.67^{*}$	$t(33) = -1.45^{°}$ $z = -1.33^{°}$

STUDY FOUR

	Mean EG	Mean CG+	Mean CG-	Mean EG-	EG+ versus CG+ (*t* test & WSRT)	EG+ versus CG- (*t* test & WSRT)
DT	5.84	5.50	5.59	5.59	*n.s.*	$t(31) = 1.250^{(p=.11)}$ $z = -1.425°$
SAT	8.22	7.79	8.14	8.14	$t(31) = 1.277^{(p=.105)}$ $z = -1.262^{(p=.103)}$	*n.s.*
SATEXP	8.33	7.97	8.36	8.36	$t(31) = 1.332^{°}$ $z = -1.373^{°}$	*n.s.*
BOOK	8.03	7.19	-	6.91	$t(310) = 1.911^{*}$ $z = -1.903^{**}$	Not available

No significant difference was found between EG+ and CG- for study 3 and no significant difference was found between EG- and both CG for study 4.

Legend:

** $p_{1-tail} < .01$; * $p_{1-tail} < .05$; ° $p_{1-tail} < .1$;
WSRT = Wilcoxon signed ranks test.
DT = DT satisfaction scale;
SAT = Oliver satisfaction scale;
SATEXP = Aurier&Evrard (1998) satisfaction scale;
SPOON = appreciation of the foldable spoon;
BOOK = appreciation f the free book.

Notes:
([a]) significant (alpha = .01) for 'real' scores (i.e., projection in a real situation): SAT: EG- = 4.23, $t(30) = -2.18$, $p_{1-tail} < .05$, $z = -1.72$, $p_{1-tail} < .05$.

TABLE 4

PEARSON (LOWER TRIANGLE) AND KENDALL (UPPER TRIANGLE) CORRELATIONS - 1 TAIL TESTS - STUDY 4

GE+ (N = 32)

	DT	SAT	SATEXP	VSUR [controlled for disc]	OBSS	FS	BS	EYE	EYEB	EXCL	EXPL
DT		.657***	.738***	.227° (p = .056) [.294*]	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
SAT	.824***		.736***	.238*[.209*]	n.s.	n.s.	n.s.	n.s.	.202°	n.s.	n.s.
SATEXP	.861***	.903***		.187°[.187°]	n.s.	n.s.	n.s.	n.s.	.224°	n.s.	n.s.
VSUR [controlled for disc]	.294*[.230°]	.250°[n.s.]	.233°[n.s.]		.230*	n.s.	.221°	.226*	.226*	-.202°	n.s.
OBSS	n.s.	n.s.	n.s.	.267°		.722***	.545***	.723***	.711***	.241°	.410**
FS	n.s.	n.s.	n.s.	n.s.	.828***		.324*	.783***	.769***	n.s.	n.s.
BS	n.s.	n.s.	n.s.	n.s.	.645***	.409**		.351*	.317*	.302*	.882***
EYE	n.s.	n.s.	n.s.	.242°	.823***	.906***	.368*		.657***	n.s.	.231°
EYEB	n.s.	n.s.	n.s.	.312*	.780***	.843***	.360*	.744***		n.s.	n.s.
EXCL	n.s.	n.s.	n.s.	-.259°	.272°	n.s.	.302*	n.s.	n.s.		.342*
EXPL	n.s.	n.s.	n.s.	n.s.	.495**	.270°	.882***	.233	n.s.	.342*	

GE⁻ (n = 32)

	VSUR	OBSS	FS	BS	EYE	EYEB	MOUTH	EXCL	EXPL	VNE	FNE
VSUR		n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	.420***	n.s.
OBSS	n.s.		.748***	.391**	.480***	.504***	.465**	.326*	.278*	n.s.	.213°
FS	n.s.	.854***		n.s.	.581***	.681***	.335*	n.s.	n.s.	-.200°	n.s.
BS	n.s.	.436**	n.s.		.293*	n.s.	n.s.	n.s.	.878***	n.s.	n.s.
EYE	n.s.	.638***	.681***	.294*		.355*	n.s.	n.s.	n.s.	-.218°	n.s.
EYEB	n.s.	.606***	.724***	n.s.	.550***		n.s.	-.220°	n.s.	n.s.	n.s.
MOUTH	n.s.	.531***	.404*	n.s.	.260°	n.s.		.670***	n.s.	n.s.	.479**
EXCL	n.s.	.330*	n.s.	n.s.	n.s.	n.s.	.599***		n.s.	n.s.	.475**
EXPL	n.s.	.278°	n.s.	.878***	n.s.	n.s.	n.s.	n.s.		n.s.	n.s.
VNE	.511***	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.		n.s.
FNE	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	.592***	.475**	n.s.	n.s.	

*** $p_{1\text{-}tail} < .001$; ** $p_{1\text{-}tail} < .01$; * $p_{1\text{-}tail} < .05$; • $p_{1\text{-}tail} < .07$. Pearson correlations for non-normal variables have to be considered with caution

FIGURE 1

THE SYNDROME OF SURPRISE (IN BLACK) AND STIMULUS EVALUATION CHECKS
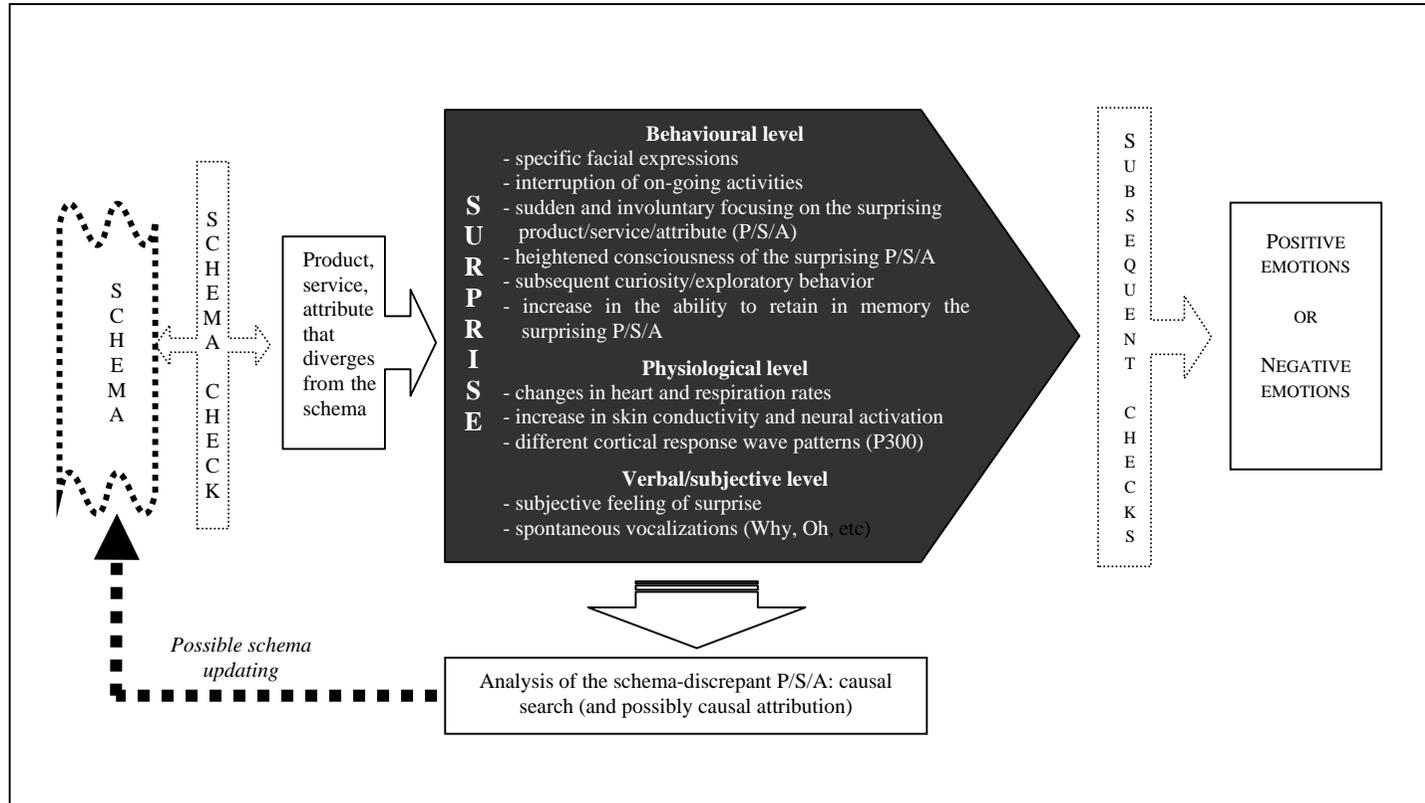
FIGURE 2

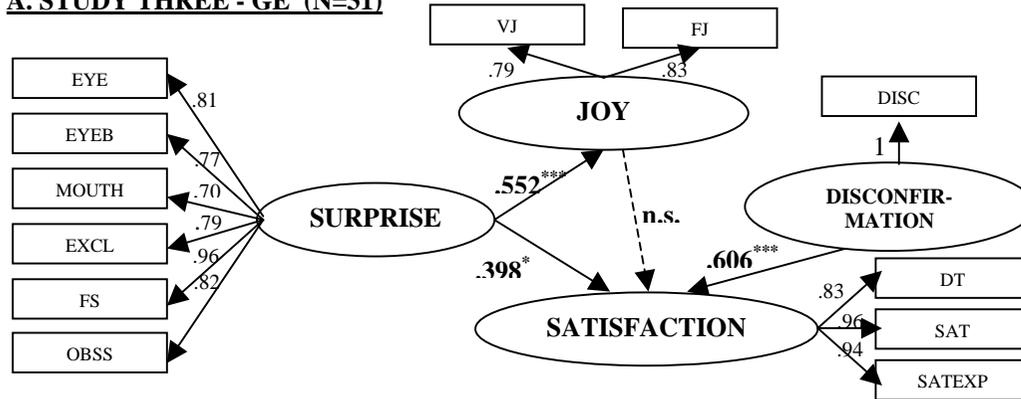THE 'SURPRISE – SATISFACTION' DIRECT AND INDIRECT RELATIONSHIPS
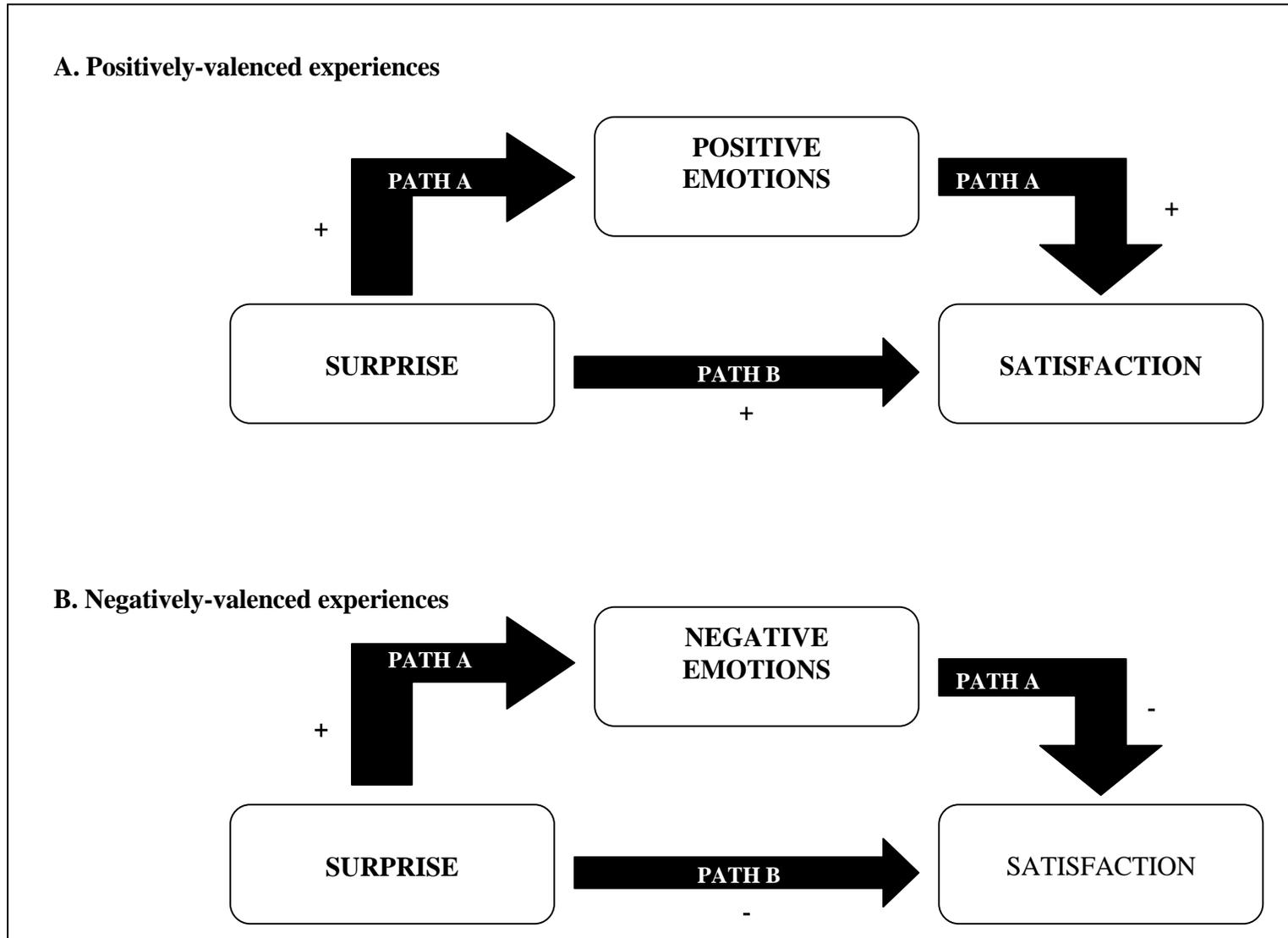
FIGURE 3

PLS APPROACH FOR MODELING STRUCTURAL EQUATIONS WITH LATENT VARIABLES (MODE A) - LVPLS 1.8 (LOHMÖLLER J.B.)
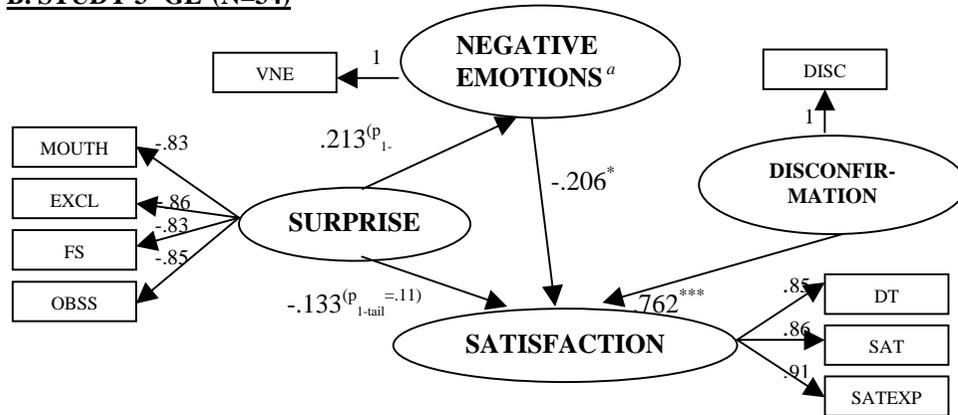
Legend figure 3:

***p<.001; **p< .01; *p< .05; •p< 0.07; 1 tail tests.
EYE=coding of opened eyes;
EYEB=coding of raised eyebrows;
MOUTH= coding of opened mouth;
FS= general measure of facial expression of surprise (taking into account EYE, EYEB and MOUTH);
EXCL= apparition of a sponteanous exclamation of surprise; STOP= interruption of ongoing activities;
EXPL= focus of the attention; OBSS=coding of general observed surprise (all aspects included);
VSUR=verbal surprise;
VJ=verbal joy;
FJ=facial expression of joy;
DISC=disconfirmation (see also table 2 for satisfaction scales);
SURPRISE, JOY, DISCONFIRMATION, SATISFACTION and NEGATIVE EMOTIONS are latent variables (their distribution is normal);

Coefficients are standardised coefficients; Distributions of latent variables are normal.

[a] No facial expression of negative emotions has been recorded by the judges.

Publications in the Report Series Research* in Management

The Effect of Members' Satisfaction with a Virtual Community on Member Participation
Fred Langerak, Peter C. Verhoef, Peeter W.J. Verlegh, Kristine de Valck
ERS-2003-004-MKT

Surprise … Surprise, An Empirical Investigation on How Surprise is Connected to Customer Satisfaction
Joëlle Vanhamme
ERS-2003-005-MKT