The Impact of Expectations on User Experience: Surprising the User

Alice Marlene Gross  
TU Berlin  
Marchstraße 23 12049 Berlin  
agross@zmms.tu-berlin.de

Juliane Bürglen  
TU Berlin  
Marchstraße 23 12049 Berlin  
jbuerglen@tu-berlin.de

Exciting product design and great user experience (UX) have become one of the most important selling points for digital interactive products. Strong competition fosters innovative user experience design and the development of new design techniques. Surprising users by using unexpected product features has proven to be beneficial for product user interaction and UX. Unfortunately, these findings have been mostly limited to classical product design. In two studies we have tried to elicit surprise in users while they were playing a digital Tetris game. Within this context we investigated the influence of pleasant and unpleasant surprises on UX ratings of the game. Both manipulations had an influence on the affective states of users, pleasant surprise elevating the users' emotional state, and vice versa. Interestingly, adding unpleasant surprise in the form of challenging messages elevated product loyalty for products associated with unpleasant surprise compared to products associated with neutral or pleasant surprise.

User Experience; Emotional Design; Surprise, Design.

1. INTRODUCTION

Since the late 1980s, user emotions have been an important part of product design [5]. Besides the technological and the inherent characteristic features of a product, emotions serve as one of the main pillars of persuading a user into liking and using a product [4]. Emotional relationships or emotional bonds thus define a particular kind of user-product relationship and are built upon a deep and sustained meaning [5]. One important result of an emotional bond with a product is the tendency to keep the product as long as possible, not replacing it with newer, nicer or better working ones [28]. All in all, it has been stated that in today's market situation it is not sufficient to design products that are merely functioning and usable, but to design products that yield an emotional experience for the user [32]. This approach is called emotional design.

1.1 EMOTIONAL DESIGN APPROACHES

One of the main tendencies in emotional design is to design for positive emotions. Fun and enjoyment are the core emotions designers strive to evoke through their product design. Blythe et al. [2] coined the term funology to describe this phenomenon. Expanding the importance of positive emotions in relation to product design, Jordan [13] promotes positive affect, namely pleasure, as a source of a flourishing user-product relationship. In his famous book about the design of pleasurable products he identifies four main sources of pleasure: physio-pleasures which are experienced through the sensory organs, socio-pleasures which are elicited by social contacts through product usage, psycho-pleasures which are generated by reaching ones goals or accomplishing a task with the help of a product and finally ideo-pleasures which are connected to the values and artistic pleasures of a product.

Desmet and Hekkert [6] propose a framework for product experience that describes three levels of affect generation: 1) The aesthetic level where all affect is caused by sensory stimulation; 2) the meaning level where all affect is induced by the expressionist characteristics of a product, and finally 3) the emotional level where affect is attributable to an object and combined with an appraisal of this object, typically resulting in an emotion.

Norman [23] identifies three levels of design, all of which are playing a crucial role in shaping a user’s experience. Visceral design refers to all aspects of a product that elicit a quick emotional response within a user without having to effortfully process it on a cognitive level. Norman further describes the behavioral level, referring to all design aspects connected to performance and usability. Finally, he proposes the reflective level which is associated with the highest amount of cognitive
Surprise constitutes a special kind of emotion. Following Russell’s two dimensional organization of emotion, it is an emotion high in arousal and neutral in valence [29]. Since the feeling of being surprised itself is not linked to a specific valence, some researchers take this lack of ‘emotional characteristic’ as proof for the fact that surprise cannot be identified as an emotion. Other authors recognize surprise as being an emotion, since it is followed by a valenced reaction most of the time, resulting in a pleasant or unpleasant surprise [1]. The discussion of all differing views and positions on the topic would go beyond the scope of this paper (see [1] and [14] for a more detailed discussion of the problem). What emotion theorists seem to have agreed upon when it comes to defining emotions and specifically surprise is to either take a categorical approach or use appraisal theories to define similarities and differences between emotions. Categorical approaches separate emotions by defining ‘basic emotions’ which are cross-cultural, universally recognizable and fundamentally different from each other like for example fear, happiness, anger or shame. They furthermore identify emotions that do not possess these universal characteristics and might be a mix of two or more ‘basic emotions, like hope or pride [7; 11; 34]. Contrary, appraisal theorists see emotions as a result of an individual’s evaluation and interpretation (e.g. appraisal) of events in a specific environment [16]. In the particular case of surprise, these appraisals have been associated with unexpectedness, novelty, unfamiliarity, amazement or pleasantness [26; 28; 30]. Following Reisenzein’s belief-desire theory of emotion (BDTE) we define the emergence of surprise as being a combination of expectations and desires about the occurrence of an event [26]. Consider for example the new smartphone you just bought. You expect it to be exceptionally easy to synchronize all your contacts as your friends told you about how smooth it was for them. Additionally, all test reports in magazines praised the outstanding ease of this process. So you open your account and you press synchronize and you are confronted with an error message, saying all your contacts have been lost. You are surprised and then you will feel some unpleasant emotions like frustration or anger. You are experiencing unpleasant surprise because your expectations were not fulfilled and your wishes were not granted, either. Now imagine the opposite situation: Your smartphone broke and a friend gives you his old crappy one for the time in between. You remember him telling you how annoying the synchronization was and how he had to put each contact into his phone by hand. So you open your account and you press synchronize and all your contacts appear in your phone book, nicely sorted and complete. First, you are surprised and then you will feel some pleasant emotions like joy or amazement. Now you are experiencing a
pleasant surprise because something unexpected happened that you did hope for (see figure 1 for a schematic description of these processes).

In this paper we try to transfer the findings from classical product design to the domain of interactive digital products.

2. RELATED WORK

To understand why and how surprise can influence the experience a user has with a product, we will give a brief overview of the temporal development of UX, as well as of some research on classical product design and surprise.

2.1 User Experience Lifecycle Model ContinUE

The User Experience Lifecycle Model ContinUE by Pohlmeyer [25] distinguishes between several use phases which can occur during the interaction with a product (see figure 2). According to ContinUE, UX already begins before the actual interaction in the so-called pre-use phase. Users build on their prior experiences with the same or a similar product as well as personal attitudes to form expectations about the interaction. During the use-phase, user, system and context form the interaction experience. When usage is over, users enter the post-use phase where they reflect their experience. Other phases of the model describe the repetitive use of a system and the resulting commitment to the same or different systems/products. When looking at the surprise formation process described in the previous section, one way to influence UX through surprise is through expectations.

2.2 Surprise as a design factor

Ludden and colleagues have conducted a very interesting line of research about surprise as a design factor in classical product design [16]. They were able to create surprising products by equipping them with visual-sensory incongruities [18]. Products like a newspaper stand, a cup or a softbox were looking like they were made out of a certain material, (fabric, stoneware, felt) but felt fundamentally different (inflexible and rough, soft or hard). Participants reported a higher overall liking, high amusement and more fun after having interacted with these products. Elevated word-of-mouth has also been reported in connection with surprising product design [17]. Some researchers have made use of the concept of pleasant surprise in connection with Wow. The concept of the Wow experience is assumed to describe an experience that exceeds the expectations of a user and leaves him in awe after the interaction [35]. Wow is a composition of several concepts, one of them being the pleasant surprise [5]. Our purpose of this article is to shed light on a possible transfer of the beneficial findings of pleasantly surprising product design to the domain of digital, interactive products. Additionally, we would like to investigate the possible negative impact of unpleasant surprise on UX ratings of digital products.

2.3 Earlier findings on UX and surprise

While pleasant surprise has been studied extensively in classical product design, only few researchers have actively explored it as a design factor for digital, interactive products. Although some studies refer to surprise-related concepts, like wow, delight or appraisal [21; 35], most research was constrained to non-interactive products. A study using the game Tetris as the experimental paradigm found significant support for
the influence negative or unpleasant surprise can have on the overall UX ratings of the said game [8]. To induce different surprises during the course of the Tetris game, one group received unexpected additional points, whereas another group suffered an unexpected loss of points. A significant effect of the factor ‘group’ was found. However, single comparisons revealed that there was no significant difference between the two groups comparing reaction times of the experimental group with the baseline. Only the differences between the minus group and the baseline group appeared to be statistically relevant. To measure the impact of surprise on UX ratings of the game, a number of rating scales was used. The results revealed that emotional valences, as well as hedonic ratings were affected by the loss or gain of points. This effect resulted from the impact of negative surprises in the group that lost points. With respect to UX, the experiment showed that an unexpected bonus in a game may not be as surprising as one might assume.

3. RESEARCH INTEREST

Our research aim is to investigate the question of how pleasant and unpleasant surprises influence the UX ratings of digital interactive products. Building on findings from classical product design where surprise as a design element had a positive effect on users, we try to transfer these findings to the product domain of digital games. Therefore, we have asked the following research questions in a first experiment: Does the frustration of expectations playing a Tetris game, thus generating pleasant and unpleasant surprise, influence the UX ratings of this game? We argue that the creation of a valenced surprise moment during the game will have an influence on the experience users have with the game. We further argue that in order to design for a great UX, pleasant surprise could be a handy tool for designers, since it shapes the user’s experience in a positive direction. Furthermore, we want to assess the possible negative effects of unpleasant surprise on the UX of a product to be able to give valid advice to designers and developers concerning the impact of this phenomenon.

Why Tetris? We would like to shed light on the effect of unexpected events, thus surprising behaviour, on the UX of a digital, interactive product. Especially products with strong competition on the market like mobile applications or websites could benefit from insights into possible beneficial and harmful effects of pleasant and unpleasant surprise, respectively. We chose Tetris as a suitable paradigm to investigate our research question for several reasons: First, to create a surprise, one has to create expectations which can later be frustrated, resulting in a possible surprise reaction. Because Tetris is a widely known game, we were not forced to artificially create expectations for the interaction with a prototype of an application or a digital product which was not yet commonly known to the majority of users. Second, Tetris is an open source game which allowed us to create our own manipulations within the game by changing the code of an existing Java version. Third, our Tetris version could be played on a desktop computer, making it easier to log data and to send markers from the PC to an EMG recorder. To gain insight into their UX ratings of the game, participants had to fill in several questionnaires after the completion of the study (see section dependent variables). Following Reisenzein [26], we use reaction times as an indicator for a surprise reaction in users: prolonged reaction times should indicate an interruption of ongoing processes. This prolongation should vanish over a repeated presentation of unexpected events due to the assimilation of the unexpected new information into the mental model of the situation [9;26], see figure 1.

4. EXPERIMENT

4.1 Method

Participants had to play a regular Tetris game while being surprised at several points in time. Tetris is a classical puzzle video game. The player is confronted with one of seven so called tetronimoes: geometric shapes each composed of four blocks which abstractly form the letters J, L, I, O, Z, S, and T. The player consecutively manipulates a theoretically infinite number of these tetronimoes on a gaming field. By moving them sideways and rotating them, the player can create horizontal lines without any gaps at the bottom of the gaming field. Those complete lines disappear upon creation, which then in turn lets all blocks on top of them fall down. Classically, players enter a higher level when a certain number of gapless lines have been cleared. We made some minor changes to our version of the game to make it suitable for the purpose of our study. To simplify terminology, we will refrain from using the name tetronimoes and will call the geometric shapes blocks, from now on.

4.2 Participants

All participants were TU Berlin students recruited either via bulletins, which were distributed throughout several buildings of the university, or via a participant server, where students could sign up
for the study. As a prerequisite for participation in the study, participants needed to have prior knowledge of Tetris, as well as a sufficient level of German. A total of 53 participants took part in the study, of which 28 were female. They were equally distributed over the three experimental groups, sex being a balanced factor. The average age of participants was 27.6 years (SD= 6.7). All participants were enrolled students at the Berlin Institute of Technology. All of them indicated to have at least a fair knowledge of Tetris. Participants did so by filling in a 6-point scale, ranging from 1=none at all, to 6=very much, asking ‘How much experience do you have (present or prior) with Tetris?’ All participants indicated at least a three or higher on the scale. Average experience was 4.25 (SD=0.95). Experience was related to knowledge of how to play the game, yet not to any valenced prior experience or personal emotional attitudes towards the game.

4.3 Independent Variables

There were two independent variables. The first independent variable was called group which had three levels. Participants were equally distributed over three groups and thus played one of three versions of a Tetris game. Group control was surprised neutrally three consecutive times by a pop-up window with three grey circles and an update about the current score with the message e.g. 'Dein Punkte-Stand: 370' ('Your score: 370'). Similarly, group bonus was surprised three consecutive times by a pop-up window with a positive smiley face and the message 'Super! Das sieht gut aus!' ('Great. This looks good!'), and group minus was surprised three consecutive times by a pop-up window with a negative smiley face and the message 'Ohje! So wird das nichts!' ('Oh no! That's not gonna work out!'). Each message had the same amount of words and visual information in it to keep the amount of information processing equal over all three groups in order not to compromise processing times and thus reaction times. The pop-up windows appeared at the outer top right corner of the Tetris game, not covering any part of the game. Additionally to the visual feedback, every group received an auditory feedback consisting of a neutral white noise (control), a crowd clapping (bonus), and a crowd booing (minus). Auditory signals were of same length (2sec). See figure 3 for examples of the pop-up windows. Group was operationalized as a between-subjects variable. The second independent variable was the time at which an unexpected event occurred. Participants were surprised by a message on the monitor during the Tetris game three consecutive times, which occurred in block 38, 47, and 51. These three different moments of surprise manifest the variable event (e.g. event1, event2, event3). This variable was operationalized as a within-subjects variable. In total, all participants had to handle 55 blocks. In contrast to a conventional Tetris games, the score was calculated by adding 10 points for each block which was displayed on the Tetris field. This allowed for the game score to serve as the discrete indicator for an unexpected event. At the same time, it could be controlled how many blocks each participant had handled before an event.

![Figure 3: Surprise pop-ups from left to right: neutral, bonus, minus.](image)

4.4 Dependent Variables

There were six dependent variables, four variables to cover UX ratings and two variables as objective measures for surprise. To measure UX ratings, the following questionnaires were used: the SAM questionnaire which is a 9-point non-verbal instrument for the evaluation of emotions, measuring the dimensions valence and arousal [3], the AttrakDiff mini, a 10-item semantic differential with the subscales pragmatic quality, hedonic quality identification, hedonic quality stimulation, and attractiveness [11], and the meCUE questionnaire. The meCUE is a validated modular questionnaire, covering many aspects of UX. The questionnaires structure is based on the Components of User Experience model (CUE model) by Mahlke and Thüring [20]. The model distinguishes between task-related and task-unrelated product qualities. Additionally, users' emotions are integrated into the model as a crucial, mediating factor for the consequences of use like acceptance or overall rating. The meCUE has been developed to give researchers an agile tool when investigating the UX of a given product. There are three modules in the meCUE: Product Perception which incorporates/ covers the subcategories utility and usability, as well as visual aesthetics, status, and bonding. The module Emotions covers positive as well as negative emotions. The module Consequences covers product loyalty, intention of use and a global rating (see [22] for a more detailed description of the development and validation of the questionnaire). For the current study, the modules user emotions and consequences of use were employed with a total of 19 items. To explicitly measure subjective surprise, a single item ranging from 1=not surprised at all, to 7=very surprised, was given to participants after completion of all other questionnaires. As in experiment 1, reaction times were used as an objective measure for surprise. To gain further insight into physical reactions to surprise, facial
EMG was conducted, measuring activity of the lateral frontalis (LF) muscle which has been associated with the facial manifestation of surprise reactions [36]. EMG data was recorded using the Heally Health Lab System by Space Bit, from Koraleski Industrie Elektronik. Data was recorded over the lateral frontalis on the right side of the face with Kendall 24mm disposable electrodes. To reduce movement and blink-related artefacts, as well as the positive skew, data were submitted to a 15-Hz high-pass filter, fully-rectified, and were then subjected to a square-root transformation. EMG reactivity was measured as the difference between activity during 1000ms before the stimuli (e.g. surprise events) and 1500ms after stimuli onset.

4.5 Hypotheses

Our general assumption is that pleasant and unpleasant surprises have differing effects on the UX ratings of the Tetris game. We chose to manipulate the Tetris game in three ways to create a pleasant surprise, an unpleasant surprise, and a neutral surprise. To verify a successful surprise manipulation, reaction times and facial EMG data were analyzed. Because all three groups experienced a surprise manipulation, a baseline was computed to compare surprising episodes with non-surprising episodes. The baseline was defined as mean reaction times of all participants averaged over the three handled blocks prior to each surprise event (e.g. block 37, 46, 51). Accordingly, this baseline was then subtracted from actual reaction times. It was assumed, according to Reisenzein [26] that these difference values of reaction times after a surprising event will be longer than after the average reaction time of the three prior blocks with no unexpected event (hypothesis one). Furthermore it was assumed in a second hypothesis that this elevation in reaction times will subside from the first to the third surprise event according to Reisenzein’s BDTE [26]. The second set of hypotheses is concerned with the effect of surprise on UX ratings. We expect a main effect for group indicating that the event in group bonus will elevate UX ratings (hypothesis three), whereas an unpleasant surprise like the message in group minus will lead to a decrease in UX ratings (hypothesis four).

5. RESULTS

To avoid an overshadowing of their final UX ratings by the negative experience of not reaching their goal, three participants who were not able to finish the game, resulting in a game over, were excluded from further analysis. Furthermore, Exclusion from EMG data analysis due to technical problems left group sizes of 14, 14, and 13 persons for bonus, minus, and control, respectively.

5.1 Surprise manipulation

5.1.1. Reaction Times

To properly investigate the effects of the unexpected events on reaction times and thus information processing, a baseline for reaction times was calculated. To do so, mean reaction times from three blocks prior to an eventful block (e.g. block 37, 46, and 51) were averaged. This baseline time was then subtracted from the individual reaction times for each block for every participant in each group. Every reaction time longer than zero milliseconds (ms) thus indicates a reaction time higher than the average reaction time during all games over all participants for three blocks prior to blocks paired with event. To verify a successful surprise manipulation, reaction times were analyzed using SPSS. Figure 4 (top) shows mean reaction times per group and per event minus the baseline reaction times (top); Mean activation over lateral frontalis per group (bottom).
difference between events1 and event2 (p=.002), event1 and event3 (p=.000), and event2 and event3 (p=.009). This indicates the typical disappearance of the surprising characteristic of an unexpected event after several occurrences [23].

There was no significant main effect for group (F(2,36)=0.195, p=.823, η\text{partial}²=.010) indicating a similar surprise reaction throughout participants in all three groups. Furthermore, there was no significant interaction between group and event (F(4,76)=.587 p=.673, η\text{partial}²=.030).

5.1.2. EMG Data
Several participants had to be excluded from EMG data analysis due to technical problems, leaving a group size of 14, 14, and 13 persons for bonus, minus, and control, respectively. EMG data was recorded and processed according to the description which can be found in the section 4.4, dependent variables. The ANOVA on the activity over lateral frontalis revealed no significant main effect for group (F(2,38)=3.00, p=.062, η\text{partial}²=.136). Furthermore, there were no significant effects (Greenhouse-Geisser corrected) for event (F(3.342, 63.501)=.244, p=.744, η\text{partial}²=.006) and event*group (F(3.342, 63.501)=.886, p=.462, η\text{partial}²=.046). See figure 4 for a descriptive picture of activities over lateral frontalis after each unexpected event.

![Figure 5: Mean SAM ratings per group.](image)

**Figure 5:** Mean SAM ratings per group.

### 5.2 Influence of surprise on UX

Mean ratings for the dimensions of the UX-questionnaires and the surprise item are shown in table 1. To investigate surprise effects on UX ratings, a one-factorial MANOVA was carried out with ‘group’ as between-subjects factor and ‘experience’ as a covariate (all values z-transformed). The MANOVA revealed a significant multivariate main effect for group (Wilks’s Λ = 0.447, F(20,78) = 1.930, p = .021, η\text{partial}²=.331). Power to detect the effect was .959. Experience did not have a significant effect on questionnaire ratings (Wilks’s Λ = 0.672, F(10,39) = 1.900, p = .075, η\text{partial}²=.328).

Given the significance of the overall test, the univariate main effects were examined. Significant univariate effects for group were obtained for both subscales of SAM, and the meCUE product loyalty scale. No effects were found on any of the AttrakDiff mini scales. Contrast analysis revealed that on the SAM Valence scale, participants in bonus rated the game significantly better (p=.018) than participants in the other groups. For SAM Arousal, control indicated lower overall arousal levels than minus (p=.006) and bonus (p=.010). On the meCUE, participants in minus indicated a higher subjective product loyalty than control (p=.009). The subjective surprise ratings showed no significant main effect for group (F(2,49=3.015, p=.058), although a slight tendency for bonus and minus being subjectively more surprising can be observed.

### 5.3 Discussion

The goal of the study was to test the influence of pleasant and unpleasant surprise on the UX ratings of a digital, interactive game. Success of the surprise manipulation was verified by analyzing reaction times following a surprising event. Support for the hypothesis concerning a prolongation of reaction times after a surprise was found conducting an ANOVA. The significant prolongation of reaction times for surprising events in all three groups, as well as the subjectively indicated levels of surprise which were marginally significant, show the success of the surprise manipulation, confirming hypothesis one. A decrease of this effect from event1 to event2 to event3, indicates the typical disappearance of the surprising characteristic of an unexpected event after several occurrences, confirming hypothesis two.

Based on these reaction times analyses, it can be assumed that surprise manipulation was
Interestingly, although users indicate worse UX, participants achieved their goal later, thereby indicating otherwise. Practically speaking, the Tetris game in group minus actively told participants that they are no good in playing the game. All participants achieved their goal later, thereby proving the games predictions to be wrong. This could have resulted in a higher liking of the game.

6. SUMMARY AND CONTRIBUTION

Eliciting positive emotions through surprise has been shown to work very well in classical product design [16; 17; 18]. In the UX community, it is largely agreed upon that a positive UX is a conglomerate of –among others- experience, design, and emotions [15]. One promising way to trigger a pleasant surprise could thus be the creation of a positive surprise reaction for users of digital products. In addition to investigating the possibility of this transfer from classical product design studies of pleasant emotions, we were interested in the severity of possible negative effects following an unpleasant surprise. Considering the influence of surprise on UX ratings, our results indicate the success of a transfer of the findings from classical surprising product design. We want to stress several points while discussing our findings. Pleasant surprise has a positive effect on UX ratings by elevating affect ratings as well as hedonic ratings of a product. Unpleasant surprise has the opposite effect on UX ratings, leaving users in a worse mood after interaction than they were before as has been shown in earlier studies. Interestingly, although users indicate worse UX ratings after they have been unpleasantly surprised, they show higher product loyalty concerning the Tetris game than users who experienced a pleasant or neutral surprise. According to findings from Fokkinga and Desmet [8] this finding can be explained in terms of frustration and satisfaction, two feelings that are part of a challenging experience. By telling participants that they will not succeed in the game, a certain frustration is build up which is later resolved into a feeling of achievement when the game has been conquered. The video game lives from the paradoxical relation between frustration and achievement. Furthermore, it gives the participant a pleasant feeling in the end, resulting in a higher product loyalty compared to the other games.

From a marketing perspective our study thus raises several questions: Since pleasant surprise has a positive impact on UX ratings, possibly increasing the market position of a product, it could be beneficial to create surprise. This could be implemented from the beginning of the use of a product by not revealing all positive features of a system and covering up some of them so that the user can experience them unexpectedly by himself. The question here is: Is it more beneficial to tell customers all positive aspects to prompt them to purchase the system? Or is it better to let them discover some surprising extras later which might pay off in the long run by increasing brand loyalty? Furthermore, unpleasant surprise seems to put a certain challenge factor into the game, which participants liked and which ultimately elevated the product loyalty. Outcomes of the study thus did not deliver clear cut implications for product designers in terms of using surprise deliberately as a design feature. Both manipulations had a positive effect on UX, the question maintains in which context which manipulation is advisable. Obviously, our results are not far-ranging enough to provide a sound answer, but future investigations may shed more light on this issue.

There are some limitations within our study: One major limitation of the study is the choice of paradigm. By choosing Tetris, we limited our research to the domain of digital games. Games are designed to enhance players emotional and affective states, so experienced gamers could have been conditioned to experiencing surprise and vice versa. Therefore, using the Tetris game as an experimental paradigm could impact on its validity: A generalization of our findings to other product domains is difficult due to the special features of games. A negative surprise which had some positive effects on product loyalty could be absolutely devastating for UX in a goal-directed context, where reaching a goal quickly and efficiently is the users first wish. The challenge factor of the Tetris game is what makes it
interesting for users; however a goal-directed task should be challenge-free to ensure easy, quick, and successful completion. To make up for issues which could possibly arise from choice of context (e.g. gaming context), a change of settings is inevitable for future studies. By extending this study using different technologies in a different domain which is not necessarily goal-directed and with more common-place technologies like mobile phones or tablets valuable insights can be gained.

7. REFERENCES


