Examining frustration and performance when priming user expectations and providing a technology malfunction

Nina Ferreri and Christopher B. Mayhorn
North Carolina State University, Raleigh, North Carolina

As digital technology develops, users create expectations for performance that may be violated when malfunctions occur. This project examined how priming expectations of technology performance (high v. low v. no) and experiences of technology malfunction (present v. not present) can influence feelings of frustration and performance on a task. A preliminary sample of 42 undergraduate participants completed a QR code scavenger hunt using the augmented reality mobile app, ARIS. Following the task, participants reported what they found for each scavenger hunt clue, their responses to failures in digital technology, and technology acceptance attitudes. Several factorial ANOVAs revealed a main effect for expectation on adaptive items of the RFDT scale and a main effect for malfunction on performance level. This suggests a potential contradiction between attitudes and behaviors when considering a common scenario involving technology.

INTRODUCTION

Digital technology is a tool for creating a more comfortable and efficient lifestyle. As it develops, users create expectations for how it should perform and often times, these expectations might be unrealistically high due to the provision of biased, marketing information provided by companies hoping to sell their products. Unfortunately, expectations are not always met when malfunctions arise, and technology does not perform as expected, inducing frustration. The expectations that users develop is an example of priming where previous information unconsciously informs a person's future behavior (McDermott, 1997). There has been much research on the effects of priming in several fields regarding how it can elicit certain behaviors, including expectations of technology performance and usefulness. Additionally, there has been examination into reactions when technology malfunctions occur, thereby contradicting user expectations. Thus, priming users for a specific expectation of technology performance may influence their feelings of frustration when they encounter technology malfunctions.

Power of Priming

Triberti, Villani, and Riva (2016) examined priming effects and applied them to a situation involving technology. Their research examined if individuals could be primed to evaluate a mobile device's usefulness based on the goal they were given to complete. These goals included either locating a hospital or contacting a friend in a virtual reality (VR) environment. Both were impossible to complete as there was no hospital or means of communication present in either VR. This promoted goal-striving at an unconscious level that lasted until the second phase where participants were asked to rate the usefulness of a mobile device's interface, as an unrelated task. The presented interfaces included either a map or a telephone book application, among controlled applications, and were suspected to be rated based on whether they could have helped participants complete the goal from phase one. The results revealed an interaction where those who were primed to locate the hospital and received the interface with a map reported the device as more useful than those who did not receive the interface with a map. These findings suggested that priming individuals to complete a specific goal can also prime them to interpret seemingly unrelated devices according to the goal they were pursuing.

Previous literature has demonstrated how priming can influence expectations about behaviors regarding technology (Triberti et al., 2016). Roscoe, Wilson, Johnson and Mayra (2017) extended these efforts by exploring how priming can influence expectations about technology. They sought to understand if priming can facilitate a specific expectation of performance for an automated writing assessment system. The priming involved instructions that described two functions of the system as either well-established or under development. To measure the effects of the conditions, participants were asked for their initial expectations of the system after being exposed to the instructions, but before encountering the system. After their interaction, participants were asked to indicate the extent to which their expectations shifted in either a positive or negative way. The results revealed that when the writing system was described as well-established, participants held more positive expectations for its performance than when it was depicted as a work-in-progress. In sum, those informed that the system was well-established were successfully primed to have positive expectations of its performance.

Malfunctions and Frustration

The aforementioned literature contributes to the interest in examining priming effects with technology. To supplement this phenomenon, frustration with technology has also been addressed. Ceaparu, Lazar, Bessière, Robinson, and Shneiderman (2004) first identified the frequency, causes, and level of severity of frustrating experiences people encounter when they interact with computers. Participants engaged in one hour of personal computer use and recorded any frustrating experiences they encountered. The most common cause of frustration reported was due to malfunctions and unexpected behaviors of the technology. This comes as no surprise as a malfunction physically interrupts task completion and influences the emotions facilitated by it, specifically frustration. This poses an empirical question considering the role of malfunctions with technology on user's experience of frustration depending on the type and severity of malfunction encountered.

Zimmerman, Sambrook, and Gore (2014) examined how a malfunction would affect performance on a task directly after experiencing a malfunction within the task. The experiment showed participants 30 images, followed by a question about the image. To assess the effect of malfunction on frustration and performance, the image was either shown for three seconds or malfunctioned and showed for only half of a second. In both conditions, the question was presented for six seconds, and participant responses recorded. Following the task, participants reported their frustration, anger, pleasure, and happiness. Results indicated that participants were angrier and more frustrated if they experienced a malfunction, compared to those who did not. Additionally, frustration was correlated with a detriment in performance on the picture task, as it was speculated that frustration influenced users to make more errors when completing the task.

Given the literature addressing priming effects on expectations and behavior (Roscoe et al., 2017; Triberti et al., 2016), and frustration due to technology malfunction (Ceaparu et al., 2004; Zimmerman et al., 2014), the current study examined the combined influence of the two factors on feelings of frustration and task performance, specifically regarding augmented reality (AR) technology. A 3 (primed expectation: high v. low v. no) x 2 (malfunction presence: yes v. no) between-subjects factorial design tested how exposure to the six conditions might elicit feelings of frustration and affect task performance when primed to have a specific expectation of the AR technology performance and presented with a malfunction or not. Three hypotheses address these speculations:

H1: Those who are primed to have low expectation of the AR technology performance will report less frustration, than those who are primed to have high and no expectation of performance.

H2: Those who experience a malfunction with the AR technology will report more frustration and exhibit lower task performance than those who do not experience a malfunction.

H3: An interaction is expected, such that those who are primed to have a high expectation of the AR technology performance and experience a malfunction will report more frustration than those who are primed to have no or low expectation of performance and also experience a malfunction.

METHOD

Participants

A preliminary sample of 42 undergraduate participants were recruited from North Carolina State University. As the study is ongoing, a full sample of 72 undergraduate participants is expected to be recruited. Currently, 90.5% of the sample is between the ages of 18 and 20, with the remaining 9.5% between 21 and 23. Approximately sixty-two percent (61.9%) identify as female.

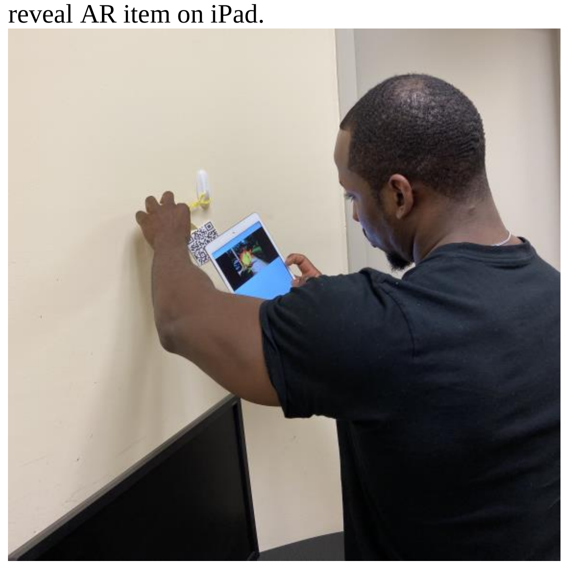
Design

To test the effects of a technology malfunction and expectation prime, participants were randomly assigned to six experimental conditions. The conditions included no malfunction present x high expectation prime; no malfunction x low expectation prime; no malfunction x no expectation prime; yes malfunction x high expectation prime; yes malfunction x low expectation prime; yes malfunction x no expectation prime.

Materials

Participants used the augmented reality mobile application, ARIS (Augmented Reality and Interactive Storytelling), on an iPad to scan QR codes hidden in a room. They were given a scavenger hunt worksheet with non-sequential clues for the physical location of the QR codes, which when scanned, revealed a gif (Figure 1). The clues were non-sequential so if a participant got stuck or missed one, then they were not prohibited from completing the task.

Figure 1. Participant using ARIS app to scan QR code and reveal AR item on iPad.



Stimuli. Stimuli included the priming expectation and malfunctioning QR code. The expectation primes were presented as scripts read by a researcher to explain the function of the ARIS app. These were similar to the descriptions presented by Roscoe et al. (2017) where three identical scripts promoted high, low, or no expectation of the performance of the ARIS app. The high expectation script primed participants to expect no problems with the application, describing the company as "seasoned" and ARIS as "their most successful app with a 4 out of 5-star rating." The low expectation script primed participants to be less hopeful regarding the app's performance, describing the company as "a start-up" and ARIS as "their first application, still in development." Finally, the no expectation script primed participants to be neutral, by simply describing ARIS as "an Augmented Reality and Interactive Storytelling app."

The malfunction stimulus was part of the scavenger hunt activity. All participants received the same set of 10 clues to guide them toward physical objects hiding 10 QR codes in the room. The QR code under the computer mouse was the chosen location for where the malfunctioning QR code was placed. For the participants in the no malfunction condition, the working QR code was placed under the computer mouse and showed an animal mouse sitting on a computer mouse when scanned (Figure 2). However, for those in the malfunction condition, the error QR code was placed under the computer mouse and provided an error message when scanned, stating "Scan successful, but…This code doesn't mean anything right now. You should come back later," thus making it impossible for participants to complete the scavenger hunt entirely. Based on the methodology reported by Triberti el at. (2016), this error message was adopted to elicit goal-striving to influence participants' reported frustration following the task. Additionally, all elements of the scavenger hunt were controlled, the only difference was when a participant was in the malfunction condition, they found an error where there should be an image of a mouse, similar to the design by Zimmerman et al. (2014), where the malfunction was either present or not.

Figure 2. Image shown for computer mouse QR code.

Measures. To measure the effects of the technology malfunctions and expectation primes, participants completed an online survey. The first section asked them to report what AR items they found for each clue of the scavenger hunt. These responses were then scored to give the participants a performance rating out of ten. The next set of items measured frustration, conceptualized as maladaptive and adaptive responses to failures in technology as measured by the Responses to Failures in Digital Technology Scale (RFDT) adapted from Hadlington and Scases (2018). This scale inquired about the likelihood to engage in adaptive or maladaptive behaviors when encountering a problem with digital technology. Responses were scored on a 5-point Likert scale where 1=very unlikely and 5=very likely. Scores for all 26 items were summed, such that higher scores indicated more maladaptive responses and lower scores indicated more adaptive responses to failures with technology. The third section assessed a Technology Acceptance Model (TAM) with five perceived usefulness items adapted from Abad, Diaz, and Vigo (2010), six perceived ease of use items adapted from Venkatesh and Davis (1996) and Abad et al. (2010), and six perceived enjoyment items adapted from Venkatesh, Thong, and Xu (2012) and Abad et al. (2010). All responses for the adapted TAM were scored on a 7-point Likert scale, where 1=strongly disagree and 7=strongly agree. Scores were summed, such that a higher score on the overall TAM measure or individual factors indicated more positive perceptions and lower scores indicated more negative perceptions of the AR technology. Lastly, ten demographics items asked participants about their experience with AR technology, hours per week of tablet, smartphone, and computer use, and personal demographics.

Procedure

Before participants arrived, the researcher verified random assignment to ensure that they received the correct stimuli. Researchers ensured the room was set up with the correct QR codes to match either the malfunction present or no malfunction present condition, and that they read the correct expectation prime script to the participant, either high, low, or no expectation. Upon arrival at the testing area, participants completed the informed consent then the researcher explained the task. They read the expectation prime and introduced the participant to the ARIS app on an iPad, demonstrating how to use it and how the scavenger hunt worked with a sample clue and hidden QR code. After the demonstration, the participant was left with the scavenger hunt clues and iPad for 10 minutes to complete the task. After 10 minutes, the researcher reentered the room to give the participant access to the survey link sent to them in the study confirmation email. They answered survey items asking about their experience with the task, what they found for each clue, responses to technology when it fails to work, and behaviors of technology acceptance. After completing the survey, participants were debriefed and given course credit.

PRELIMINARY RESULTS

Descriptive statistics for the independent variables, attitudinal and behavioral scales, and demographic items are shown in Table 1. Raw means are displayed to show that all variables were normally distributed. Interestingly, 57.1% (n=24) of participants reported that they were familiar with AR technology. Yet, of that percentage, only 35.7% reported having a little experience. This may suggest that while people are familiar with AR technology, it is still fairly novel, and most do not have much experience with it. Regarding performance, individual scores were calculated by counting the number of items reported for each clue correctly. Scores were out of 10, with a mean of 4.60, suggesting that the scavenger hunt was not too easy nor too difficult. A median split was then performed around the true median, 3.00, to separate participants into low and high-performance groups. Those with a score of 0-3 were categorized as low performance (n=22), while those with a score of 4-10 were categorized as high performance (n=20).

Table 1. Summary of descriptive statistics for IVs, DVs, and demographics.

Variable	*N*	*M* (*SD*)	Median	Mode
Malfunction IV	42	1.60 (.50)	2.00	2.00
Expectation IV	42	1.93 (.81)	2.00	1.00
Familiar with AR tech	42	1.43 (.50)	1.00	1.00
Experience with AR tech	22	2.27 (.63)	2.00	2.00
Performance Score	42	4.60 (3.94)	3.00	1.00
Performance Level	42	1.48 (.51)	1.00	1.00
RFDT Scale	40	63.28 (8.81)	62.00	61.00
PU Scale	42	22.81 (4.16)	22.50	21.00
PEOU Scale	42	36.10 (3.37)	37.00	37.00
PE Scale	41	31.66 (5.39)	31.00	29.00
TAM Scale	41	90.68 (10.15)	91.00	94.00
Age	42	2.10 (.30)	2.00	2.00

Malfunction was scored as 1=present, and 2=not present, while expectation was scored as 1=high, 2=low, 3=no. Familiarity with AR tech was binary, in which, 1=yes, and 2=no, while experience with AR tech was scored on a 4-point Likert scale, in which 1=not at all and 4=a lot. As mentioned in the method, the RFDT scale was scored on a 5-point Likert scale, where 1 = very unlikely and 5 = very likely, with nine adaptive items reverse scored. The TAM scale was scored on a 7-point Likert scale, where 1 = strongly disagree and 7 = strongly agree. Lastly, age was categorical with those between the ages 18-20=1, and 21-23=2.

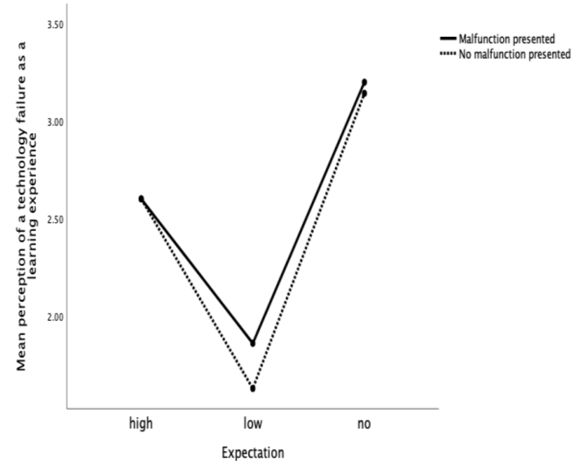
A series of factorial analyses of variance (ANOVA) were performed to determine if there were differences between the independent variables and the attitudinal and behavioral scales. Though there were no significant ANOVAs for either of the IVs with the attitudinal scales as a whole, there were several individual items of the RFDT scale that were significant for the expectation manipulation. This approach was used to examine if there were still main effects for the individual scale items that could inform continuation of this study in full. First, there were significant expectation group differences for attitudes to search for a solution to a failure in digital technology, $F(2,36)=6.12$, $p=.005$, $\eta^2=.25$. This indicated that 25% of the variability in attitudes to search for a solution was explained by the expectation group manipulation. Fisher's Least Significant Difference (LSD) comparison was used to determine which of the three expectation groups were different from one another and revealed that those in the no expectation group ($M=2.42$) were less likely to search for a solution compared to those in the high ($M=1.47$) and low expectation groups ($M=1.33$). Because this item proposed an adaptive behavior toward technology failures, the scoring was reversed such that a lower score indicated a higher likelihood of performing such behavior.

Another ANOVA revealed a significant main effect for expectation on the attitude to try everything to fix a failure with digital technology, $F(2,36)=4.11$, $p=.025$, $\eta^2=.19$. This suggested that 19% of the variability in attitudes to try everything to fix a technology failure was attributed to the expectation manipulation. A post-hoc analysis using Fisher's LSD comparison found that those in the high expectation group ($M=1.40$) were more likely to try everything to fix the problem than those in the no expectation group ($M=2.50$). This item also depicted adaptive behavior and was reverse scored as well.

A third ANOVA examined the differences between expectation group and attitudes that relish the opportunity to solve the issue. Results found a significant main effect of expectation, $F(2,36)=6.34$, $p=.004$, $\eta^2=.26$, such that 26% of the variability in responses that relish the opportunity to solve the issue were explained by the expectation level to which they were exposed. Fisher's LSD post-hoc comparison indicated that those in the high (M = 2.47) and low (M = 2.73) expectation groups were more likely to relish the opportunity to solve a tech issue, compared to those not exposed to an expectation prime (M = 3.92). This item was again representative of adaptive behavior and was reverse scored.

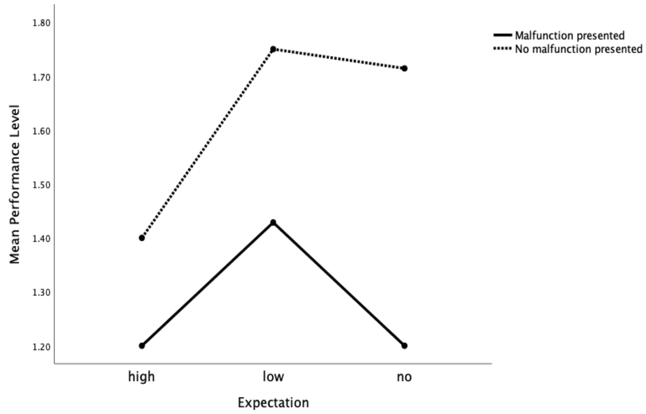
The last ANOVA regarding attitudinal scale items assessed the expectation group differences for responses that considered a tech failure a learning experience (Figure 3). There were significant expectation group differences, $F(2,36)=5.95$, $p=.006$, $\eta^2=.25$, such that 25% of the variability in attitudes that considered a tech failure a learning experience was attributed to the expectation group. The post-hoc Fisher's LSD comparison revealed that those in the low expectation group (M=1.73) were more likely to see a situation with a tech failure as a learning experience, compared to those in the high (M=2.60) and no (M=3.17) expectation groups. As this was an adaptive behavior item, the scores were reversed as well.

Figure 3. Main effect of expectation on RFDT item "consider technology failure as learning experience.



Regarding the malfunction manipulation, an ANOVA was conducted to determine if there were significant malfunction group differences for performance level. Results revealed a significant main effect of malfunction on performance level, $F(1,36)=4.91$, $p=.033$, $\eta^2=.12$, such that those who did not receive a malfunction ($M=1.60$) scored better on the scavenger hunt than those who did receive a malfunction ($M=1.29$). The malfunction groups participants were assigned to explained 12% of the variability in performance level.

Figure 4. Main effect of malfunction on performance level.



DISCUSSION

Findings from preliminary data provide support for the hypotheses. Hypothesis 1 is addressed by the results regarding the main effects for expectation on the RFDT scale items. Specifically, the findings suggest individuals exposed to the low expectation prime were more likely to see a technology failure as a learning experience and this supports the hypothesis that the low expectation group would report less frustration than those in the high and no expectation groups. The understanding is that if individuals were less frustrated, they would be more willing to resolve failures in technology. Overall these findings are consistent with Triberti et al. (2016) who explained that when a task is made impossible to complete, individuals exhibit goal-striving behavior. In the current study, such goal-striving behavior is exhibited because we observed significant results on only the adaptive attitudes for the RFDT scale.

Regarding Hypothesis 2, the significant main effect for malfunction on performance level provides support that those who do not experience a malfunction, express less frustration. Previous literature has linked frustration to performance in a negative way, such that those who were frustrated, exhibited worse performance on a task (Zimmerman et al., 2014). Similarly, the same trend seems to be present in the preliminary results of this study.

Unfortunately, Hypothesis 3 was not supported by the results. However, the presence of main effects for both malfunction and expectation in the same study indicates the importance of examining the two factors together, especially since the literature has not examined such interaction effects. There also seems to be an interesting contradiction occurring regarding malfunction and expectation, such that malfunction is linked to behavior (i.e. performance) and expectation is linked to attitude (i.e. RFDT scale items). This then poses the question of how attitudes and behaviors differ when considering a common scenario that involves technology.

Limitations. There are limitations to this study that may affect the reliability of the results. For instance, there were some known instances of unanticipated malfunctions from the ARIS app sporadically throughout the study. There were also occasional reports of the app not registering any of the QR codes, or QR codes that were not intended to malfunction. This poses an issue if the unforeseen issues occurred for individuals that were not assigned to the malfunction group. Luckily, these instances were reported very infrequently.

Future Research. As this research is still ongoing, we hope to replicate the preliminary findings with the completed dataset. For further directions, this research could continue examination into the potential dichotomy of attitudes versus behaviors. As users may be quite diverse, there is also potential to explore age effects (among other demographic considerations) that exist, as this research recruited undergraduate participants only.

CONCLUSION

The culmination of this research will provide further insight into human interactions with technology, specifically when primed to have a certain expectation of performance. It will also provide further exploration into priming effects in more technology-focused contexts, such as when technology malfunctions. This information can be particularly useful for user experience researchers when considering how users interact with products and how to better design those products to fit user needs. In this instance, user needs are related to the primed expectations, whether the company itself or societal dictations create them. Overall, understanding what factors prime user expectations to develop, and the needs they have as a result, can influence how to address product design accordingly.

REFERENCES

Abad, M., Díaz, I., & Vigo, M. (2010). Acceptance of mobile technology in hedonic scenarios. *BCS HCI.*

Ceaparu, I., Lazar, J., Bessière, K., Robinson, J., & Shneiderman, B. (2004). Determining causes and severity of end-user frustration. International Journal of Human-Computer Interaction, 17(3), 333–356.

Hadlington, L., & Scase, M. O. (2018). End-user frustrations and failures in digital technology: Exploring the role of fear of missing out, internet addiction and personality. *Heliyon, 4*(11), e00872.

Hirshfield, L. M., Bobko, P., Barelka, A., Hirshfield, S. H., Farrington, M. T., Gulbronson, S., & Paverman, D. (2014). Using noninvasive brain measurement to explore the psychological effects of computer malfunctions on users during human-computer interactions. Advances in Human-Computer Interaction, 2014.

McDermott, K.B. (1997). Priming on perceptual implicit memory tests can be achieved through presentation of associates. *Psychonomic Bulletin & Review* 4**,** 582–586.

Roscoe, R. D., Wilson, J., Johnson, A. C., & Mayra, C. R. (2017). *Presentation, expectations, and experience: Sources of student perceptions of automated writing evaluation*.

Triberti, S., Villani, D., & Riva, G. (2016). Unconscious goal pursuit primes attitudes towards technology usage: A virtual reality experiment. *Computers In Human Behavior*, *64*163-172.

van Volkom, M., Stapley, J. C., & Malter, J. (2013). Use and Perception of Technology: Sex and Generational Differences in a Community Sample. *Educational Gerontology*, *39*(10), 729-740.

Venkatesh, V., & Davis, F. D. (1996). A model of the antecedents of perceived ease of use: Development and test*. *Decision Sciences*, *27*(3), 451-481.

Venkatesh, V., Thong, J. Y. L., & Xu. X., (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly, 36*(1), 157-178.

Zimmerman, N. K., Sambrook, E., & Gore, J. S. (2014). The effects of a computer malfunction on subsequent task performance. *Behaviour & Information Technology*, *33*(9), 874-881.