The Impact of Causal Attributions on System Evaluation in Usability Tests

Adelka Niels
University of Applied Sciences
Luebeck, Germany
adelka.niels@fh-luebeck.de

Sascha R. Guczka
University of Hamburg
Hamburg, Germany
sascha.guczka@uni-hamburg.de

Monique Janneck
University of Applied Sciences
Luebeck, Germany
monique.janneck@fh-luebeck.de

ABSTRACT
Causal Attribution research deals with the explanations people find in situations of success and failure for *why* things happened the way they did, and the extent of *control* they feel to have over the situation. Attributing success and failure differently has an impact on our emotions, our motivation, and behavior. However, so far research on computer-related attributions has not answered the question whether different attribution patterns influence system evaluation in usability tests. This question formed the basis for our investigation. Two standardized questionnaires were used to measure users' attribution patterns and users' system evaluations. The usability tests were conducted in our laboratory with N=51 participants. At large, our results suggest that there are notable influences of users' attribution patterns on their evaluation of system quality, especially in situations of success.

Author Keywords
Attribution Research; Usability Evaluation; User Experience; User Types

ACM Classification Keywords
J.4. Social and Behavioral Sciences: Psychology

INTRODUCTION: ATTRIBUTION THEORY
Attribution Theory has its origin in social psychology and serves to examine, explain, and predict behavior and experiences various life domains. E.g., attributions have been extensively researched in clinical psychology, especially regarding anxiety disorders and depression.

In simple terms, the central focus of attribution research is on how people explain *why* things happened the way they did and how much control they perceive to have to influence the cause of events [10].

Attributional Dimensions
Attributional research puts special emphasis on the *effects* (consequences) of a person's attributions on his/her experiences, motivation and behavior [29]. In this regard, the four attributional dimensions of *Locus*, *Stability*, *Controllability*, and *Globality* play a crucial role [2, 28]:

Locus of control (internal vs. external) describes whether a person sees internal (e.g. 'I am to blame for this computer breakdown') or external (e.g. 'The system is to blame for breakdown') causes for an event, which influences emotional reactions. For instance, in situations of failure internal attributions are often associated with self-focused negative emotions, such as guilt and shame. Contrary, external attributions are likely to be associated with negative emotions, like anger or even aggression, which are geared towards other people or external entities [28].

The *Stability* dimension (temporally instable vs. stable) captures whether causes are perceived to change over time or not, thus affecting individuals' future expectations [17]. For example, causal factors such as intelligence or luck are generally considered as stable because they are difficult or impossible to change. On the other hand, unstable factors, such as the effort invested, are easier to change.

The *Controllability* dimension (high control perception vs. low control perception) distinguishes controllable causes from causes perceived as uncontrollable. For instance, emotions of anger, pity, guilt and gratitude might be related to the dimension of *Controllability* [28].

The *Globality* dimension (specific vs. global) describes if the cause relates to a specific subject (e.g. 'I don't understand this specific computer application') or to a global cause (e.g. 'It doesn't matter which computer application I use, I never do well).

Attributing a situation of failure within the internal/stable dimensions can lead to shame or humiliation because causes are attributed to the self and seen as unchangeable. Contrary, internal/instable attributions might also cause self-doubts and self-reproach, but the situation is seen as a singular event that will not necessarily occur again. If a situation of failure is attributed within the external/stable dimension there is less motivation to change [2].

A tendency to consistently attribute success and failure in different situations and contexts to a specific type of cause is called *Attribution Style*. Attribution Styles are considered as part of one's self-concept, which represents all of a person's self-referred attitudes [20]. Therefore, Attribution Styles can be seen as rather stable over time.

Attributional Theory in HCI

Attributional theories exist in numerous areas of psychology: In the field of motivation and emotion research [28], in clinical psychology, e.g. in the exploration of depressive disorders [2, 3], in educational psychology, e.g. regarding learning and achievement motivation [8] or health psychology [26]. The application of attribution theory has also stimulated researchers from other disciplines, e.g. organizational behavior and marketing [19, 9]. But even though there is evidence that Attribution Styles are domain-specific, the application to computer use is still a novel field in HCI research. Nevertheless, several studies have shown that Attribution Styles also exist with respect to computer-related behavior [6, 15, 16, 27].

Building on this research, we assume that different computer-related Attribution Styles have distinct influences on user experience and behavior. For example, users with different Attribution Styles might come up with quite different explanations for events like system failures, triggering different user responses. Thus, having favorable or unfavorable Attribution Styles, respectively, might account for differences regarding computer mastery, computer anxiety, or simply different styles of using computers. A detailed knowledge of computer-related Attribution Styles might help to better understand user behavior and difficulties they have when using computers. Thus, design principles can be developed to support different types of users in a specific way.

Typology of Computer-Related Attribution Styles

A typology of six central computer-related Attribution Styles was developed by Niels and Janneck [22]. The styles relate to situations of success and failure, respectively, as is commonly distinguished in attribution research [cf. 7, 24, 13]. Similar to attribution research in clinical psychology, favorable as well as unfavorable styles emerged. In the following paragraphs, stereotypical names and exemplary statements are used to illustrate the kind of attitude and behavior that might be associated with the respective Attribution Style:

Attribution Styles for success situations

- *Confident* – "I am competent and responsible for my success". Persons with this Attribution Style experience high controllability and see internal reasons for their success.
- *Realistic* persons – "Sometimes I am successful, sometimes not" – expect causes to change over time and in different situations. These persons have medium values in all dimensions except for *Stability*.
- *Humble* – "This time I was lucky". These persons experience low controllability and attribute success mainly to external causes.

Attribution Styles for failure situations

- *Confident* – "I know it was my fault, but next time I will do better". These persons experience high controllability, see mainly internal reasons for their failure, and believe causes will change over time and in different situations.
- *Realistic* – "This time I failed, but don't worry about it". These users expect causes to change over time and in different situations.
- *Resigned* – "I never understand what computers do". They experience low controllability and expect causes of failure to persist over time and in different situations. This is the most unfavorable style that can be compared to the so-called pattern of "learned helplessness" observed in patients suffering from depression [cf. 2].

Measuring Attributions and User Experience

Attribution Questionnaire

Guczka and Janneck developed a standardized *Attribution Questionnaire* (AQ) to determine Attribution Styles [12]. The AQ contains four questions to measure the attributional dimensions of *Locus*, *Stability*, *Controllability* and *Globality* and is based on the Sport Attributional Style Scale, SASS [13]. Figure 1 shows the English version of the questionnaire relating to situations of failure (items measuring attributions of success are worded analogously).

What caused the breakdown?		
1. I would locate the cause of the breakdown…		
internally (I am to blame)	1 2 3 4 5 6 7	externally (the system is to blame)
2. The cause of this breakdown is…		
a singular event	1 2 3 4 5 6 7	recurring
3. The cause of the breakdown is…		
controllable	1 2 3 4 5 6 7	uncontrollable
4. The cause of this breakdown is likely to promote other breakdowns…		
just in this situation	1 2 3 4 5 6 7	in other situations as well

Figure 1 Excerpt from the Attribution Questionnaire for failure situations

Scales	Description	Items
Attractiveness	Overall impression of the product. Do users like or dislike is?	annoying / enjoyable, good / bad, unlikable / pleasing, unpleasant / pleas-ant, attractive / unattractive, friendly / unfriendly
Perspicuity	Is it easy to get familiar with the product?	not understandable / understandable, easy to learn / difficult to learn, complicated / easy, clear / confusing.
Efficiency	Can users solve their tasks with the product without unnecessary effort?	fast / slow, inefficient / efficient, impractical / practical, organized / cluttered.
Dependability	Does the user feel in control of the interaction?	unpredictable / predictable, obstructive / supportive, secure / not secure, meets expectations / does not meet expectations
Stimulation	Is it exciting and motivating to use the product?	valuable / inferior, boring / exiting, not interesting / interesting, motivating / demotivating
Novelty	Is the product innovative and creative?	creative / dull, inventive / conventional, usual / leading edge, conservative / innovative

Table 1 User Experience Questionnaire scales and items [25]

User Experience Questionnaire

The *User Experience Questionnaire* (UEQ) developed by Laugwitz and colleagues is an established questionnaire to measure user experience – especially the perceived attractiveness, design and use quality of software systems – in a simple way [18].

The UEQ includes 6 scales with 26 items to measure the six factors *Attractiveness*, *Perspicuity*, *Efficiency*, *Dependability*, *Stimulation*, and *Novelty*. *Attractiveness* is a pure valence dimension. *Perspicuity*, *Efficiency* and *Dependability* mark pragmatic quality aspects (goal-directed), while *Stimulation* and *Novelty* describe hedonic quality aspects (not goal-directed). The questionnaire consists of word pairs of contrasting attributes that may apply to the tested system or software, respectively (Table 1). The items have the format of a seven-stage semantic differential [25].

RESEARCH QUESTIONS AND METHODOLOGY

This work touches on the research gap identified by Janneck and Guczka [15], who raise the question whether computer-related attributions have an impact on system evaluation in usability tests. Janneck and colleagues state that including computer-related attributions or Attribution Styles as personal traits in usability studies might help to better understand and interpret results. They argue that a positive or negative system evaluation could be related to attributional dimensions (of *Locus*, *Stability*, *Controllability*, and *Globality*) or to specific Attribution Styles.

Therefore, the main research objective in this paper is to investigate the correlations between different attribution patterns and system evaluation in usability tests. For instance, it is plausible to assume that people with low control perception will rate a system i.e. as less dependable, less efficient or less comprehensible than persons with high control perception, since they generally feel more in control when dealing with computers, irrespective of the specific design of the system or device.

Similarly, system quality aspects might be assessed differently depending on whether a user shows an internal (i.e., ascribing success or failure to oneself) or an external (i.e., attributing success or failure to the system) attributional pattern. For example, in situations of success external attributions may result in better user ratings on the UEQ scales.

Furthermore, users with high values for *Stability* or *Globality* might transfer prior experiences with other systems to their current usage to a larger extent, which is also likely to be reflected in their system evaluation. For example, in situations of failure high values for *Stability* or *Globality* might lead to lower ratings concerning the *Attractiveness* scale of the UEQ.

Moreover, the *Locus* dimension might influence user ratings concerning the hedonic aspects of *Stimulation* and *Novelty*. For instance, attributing success externally might lead to higher ratings regarding the *Stimulation* and *Novelty* scales of the UEQ, as users with external attribution patterns might perceive the system as more novel and thus also more innovative and interesting.

Regarding system evaluation and Attribution Styles, we assume that having favorable or unfavorable styles, respectively, might account for different user ratings. For instance, persons with a *Confident* style – who see mainly internal reasons for their failure and experience high controllability – might evaluate a system more positively in a situation of failure (because they blame themselves for the failure). On the other hand, persons with the unfavorable *Resigned* style – who experience low controllability and attribute failure mainly to external causes – are likely to blame the system for their failure and thus assess it more negatively.

Based on these considerations we formulate the following hypotheses:

H1: The *Locus* dimension has an impact on system evaluations: System quality aspects are assessed differently depending on whether users show an internal or an external attribution pattern.

H2: The *Controllability* dimension has an impact on system evaluation: Persons with low control perception will rate a system more negatively than persons with high control perception.

H3: The *Stability* and *Globality* dimensions have an impact on system evaluations: In situations of failure, persons with high values on these dimensions give lower ratings on the UEQ scales.

H4: Individual Attribution Styles have an impact on system evaluation: Persons with more favorable Attribution Styles (e.g. *Confident*) evaluate systems differently than persons with unfavorable styles (e.g. *Resigned* or *Humble*).

To figure out whether attributional dimensions or specific Attribution Styles have an impact on user experience we investigated computer-related causal attributions as well as system evaluations as part of different usability tests.

Procedure

The participants were asked to complete certain tasks using a specified application. We purposely used various applications on different devices with different tasks to ensure that possible effects occurred independently of specific system characteristics. Therefore, we were not interested in the system evaluations as such, but how they were related to attribution patterns.

The systems we tested included a traveling website, university homepage, weather forecast app and different small gaming applications that had been developed by students from a programming class. About 3-5 typical tasks were devised for each application (e.g. booking a railway ticket, investigating details of a certain study program, looking up weather conditions, game playing tasks). Most tests were conducted on a desktop computer. Mobile applications were tested with tablets. One of the gaming applications ran on a multi-touch table. Each application was tested with up to 10 different participants. None of the participants tested more than one system.

The tests were conducted in a usability laboratory. Two primary experimenters controlled and observed the tests. After completing the tests, the participants filled out the attribution and UEQ questionnaires. It was up to the participants to decide whether and – if so – how many successes or failures had taken place in the usability test (up to 3 each could be reported), as the subjective interpretation of the situation is crucial for attribution processes.

Each test session was conducted as follows:

- Participants were greeted, guided to the test station and the experimenters introduced themselves.
- The scenario for the test was briefly described to the participants by the experimenter.
- Before the participants processed the task they filled out a basic demographic data questionnaire (age, gender, educational background) and a questionnaire containing questions about their general computer use habits and experience (computer experience in years, and daily computer use in private and workspace situations, self-assessed computer skills).
- Then the participants were given an overview of the test and they were asked to solve the tasks. The participants were instructed to think aloud while working on the tasks. The experimenters observed and captured problems that occurred during the test. The total time required by participants to solve the tasks varied between 5.85 and 23.05 minutes (M=12.95, SD=4.48).
- After finishing the tasks, the participants filled out the UEQ (as described before) to rate the perceived attractiveness, design and use quality of the tested software or device.
- Then participants were asked whether success had occurred during the test and – if so – to briefly describe the cause of success, to rate the importance of the success on a Likert scale ranging from 1 (unimportant) to 7 (very important) and to fill out the AQ (as described before).
- Likewise, the participants were asked whether failures occurred during the test and – if so – how many, to rate the importance of these situations and fill out the AQ. Again, up to three failure situations could be reported.

Sample

Most of the participants were computer science students from two universities located in Northern Germany. A few of them were faculty members as well as employees from external organizations. They were not paid for their participation.

In all, 51 persons participated in the study (37% female, 61% male). Mean age was 27.96 years (range: 19-64 years). The general level of education was quite high (82% with high school or university degree). Participants were rather experienced computer users. On average they had 14.5 years (range: 7-22 years) of experience in computer use and they used computers on average 10.5 hours a day (range: 3-18 hours). Participants self-rated their computer skills on a Likert scale ranging from 1 (low) to 7 (expert) on average at 5.5 (range: 1-7).

DATA ANALYSIS AND RESULTS

For analysis we included all datasets that contained at least one complete AQ for situations of success and failure, respectively, as well as a complete UEQ. In order to examine the impact of attributions on system evaluations, we first analyzed the Attribution Styles for situations of success and failure. Secondly, we analyzed the UEQ values. Finally, we calculated correlations between UEQ scales and attributional dimensions as well as correlations between UEQ scales and Attribution Styles. In the following sections methods and results are explained.

Attribution Questionnaire

In our analysis we distinguished between situations of success and failure, as is usually done in attribution research. First, to prove the construct validity of the research instrument, Spearman correlation coefficients were calculated regarding the four attributional dimensions (Spearman's Rho was chosen due to the ordinal data quality).

Afterwards, the data was analyzed by means of hierarchical cluster analysis as an exploratory method for discovering structures in raw data [1, 4]. Firstly, we measured each subject's level of attribution per dimension. Secondly, we built a matrix, containing the distance (calculated via Euclidian measures) between the subjects regarding each dimension. After that we clustered each subject or group of subjects together, while keeping the inner cluster variance low, using Ward's method for computing the cluster linkage criterion. To rule out which cluster solution stands out, we considered the variance changes and the plotted structure (dendogram) for each data set [11].

Finally, to make Attribution Styles visible clusters were displayed as line diagrams (Figure 2 and Figure 3): High values for *Locus* mean that a person attributes reasons for success or failure to external causes. High values for *Stability* and *Globality* indicate that the cause is perceived as stable over time, and refers to a global cause; whereby, due to the wording of the questionnaire, high values for *Controllability* indicate low perception of control.

Results: Attribution Styles for success situations

Regarding inter-correlations, merely *Stability* and *Globality* correlate slightly at $r=0.35$ ($p=0.03638$). However, this is theoretically plausible: If people believe that success will persist over time they typically also believe that similar situations take place in different contexts. Thus, the construct validity of the research instrument is supported.

For success situations, cluster analysis identified three clusters, namely *Confident*, *Realistic* and *Humble* (Figure 2). Persons with a *Confident* Attribution Style usually experience high controllability and see internal reasons for their success. In this case the internality is slightly less pronounced compared to prior studies, but nevertheless this style corresponds to the *Confident* style. Persons with a *Realistic* style expect causes to change over time and in different situations. They have medium values in all dimensions except for *Stability*. Persons with a *Humble* style experience low controllability and attribute success mainly to external causes. Thus, the Attribution Styles we identified in our study were almost identical with the styles found in prior studies [15].

Table 2 shows the mean values for the clusters. ANOVAs were calculated showing significant differences between clusters. Effect sizes (according to Cohen's classification of η^2, [5]) are high.

Results: Attribution Styles for failure situations

Regarding inter-correlations, *Locus* and *Controllability* correlate significantly at $r=0.502$ ($p < 0.001$). Again, this is theoretically plausible: If people see internal causes for a situation they typically also experience higher controllability. Just as in situations of success, *Stability* and *Globality* correlate significantly at $r=0.597$ ($p < 0.001$).

For failure situations, cluster analysis identified three clusters: *Confident*, *Realistic* and *Resigned,* thus matching the attribution patterns observed in prior studies (Figure 3). Persons with a *Confident* style experience high controllability and see mainly internal reasons for their failure. Persons with a *Realistic* style expect causes to change over time and in different situations and have medium values regarding *Locus* and *Controllability*. Persons with a *Resigned* style experience low controllability and expect causes of failure to persist over time and in different situations. Compared to previous studies [cf. 22], the *Resigned* style in our study is characterized by lower values regarding all four dimensions. Nonetheless, it shows the typical characteristics of this attributional pattern.

Again, ANOVAs were calculated showing highly significant differences between clusters with high effect sizes (according to Cohen's classification of η^2, [5], see Table 2).

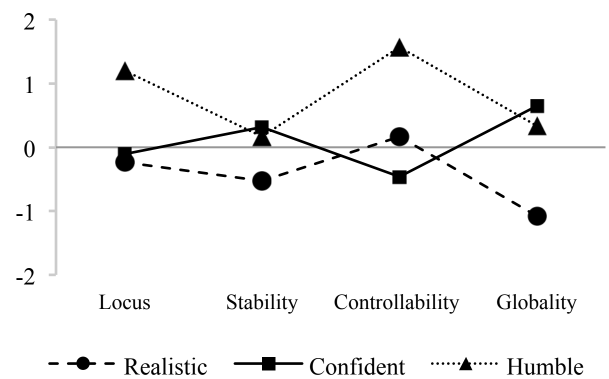

Figure 2 Standardized cluster results for situations of success

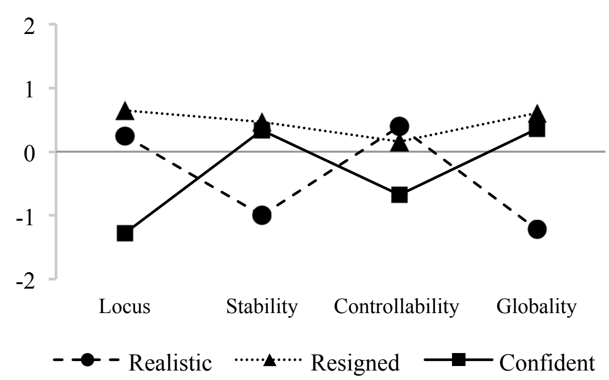

Figure 3 Standardized cluster results for situations of failure

Success Cluster	Realistic n = 18	Confident n = 27	Humble n = 6	F-value	p	η^2
Locus	3.00	3.22	5.50	5.930	0.005	0.198
Stability	5.11	6.19	6.00	4.389	0.018	0.155
Controllability	3.50	2.37	6.00	17.391	0.000	0.420
Globality	3.33	6.19	5.67	45.505	0.000	0.655
Failure Cluster	**Realist n = 15**	**Resigned n = 22**	**Confident n = 14**	**F-value**	**p**	η^2
Locus	5.40	6.32	1.86	47.241	0.000	0.663
Stability	3.87	6.36	6.14	18.020	0.000	0.429
Controllability	4.80	4.27	2.43	5.454	0.007	0.185
Globality	1.73	5.59	5.07	43.587	0.000	0.645

Table 2 ANOVA results for success and failure cluster

User Experience Questionnaire

The *User Experience Questionnaire* data was analyzed using the Excel *Data Analysis Tool* (www.ueq-online.org) provided by Schrepp and colleagues [25]. The first step of analysis is to transform the item values as the order of positive and negative terms is randomized in the questionnaire. The resulting unified scale ranges from +3 to -3. Mean values were calculated for each scale. Values <–0.8 imply a negative evaluation, values between -0.8 and 0.8 correspond to a neutral evaluation, while values >0.8 represent a positive evaluation [14].

The results show a positive evaluation for the UEQ scales *Attractiveness*, *Perspicuity*, *Efficiency*, and *Dependability*, as well as a neutral evaluation for the scales *Stimulation*, and *Novelty*. However, as we were not interested in system evaluations as such but rather their associations with attribution styles, the UEQ values are not reported here.

Relations between Attribution Patterns and System Evaluation

To analyze relations between attributional dimensions (*Locus*, *Stability*, *Controllability*, and *Globality*) and UEQ scales (*Attractiveness*, *Perspicuity*, *Dependability*, *Efficiency*, *Stimulation*, and *Novelty*), correlations (Sperman's Rho) were calculated.

Furthermore, persons with higher and lower expression of the attributional dimensions, respectively, were compared regarding their UEQ ratings by means of Mann-Whitney-U-Tests.

Finally, Attribution Styles and corresponding UEQ scores were tested globally for differences followed by post-hoc tests for pairwise comparison. Because of partially non-normally distributed data the Kruskal-Wallis-Test was used instead of analyses of variance.

Correlations between Attributional Dimensions and UEQ Scales in situations of success and failure

In *situations of success*, data analysis revealed positive significant correlations between *Locus* and *Perspicuity* (goal-directed quality) as well as *Novelty* (hedonic quality). Users with high values at the *Locus* dimension see external reasons as the cause for their success, thus believing that success relates to system design rather than their own computer skills. For the other dimensions no correlations were observed (Table 3).

In *situations of failure*, analysis showed a significant positive correlation between *Stability* and *Dependability*. Thus, users who perceive causes as stable over time (high values on the *Stability* dimension) evaluate systems as more reliable than persons with low *Stability* values. For the other dimensions no correlations were observed (Table 3).

Success	UEQ Attractiveness	UEQ Perspicuity	UEQ Dependability	UEQ Efficiency	UEQ Stimulation	UEQ Novelty
Locus	0.243	**0.330***	0.035	0.207	0.263	**0.500****
Stability	0.146	0.202	0.190	0.216	0.138	0.009
Controllability	0.049	0.035	–0.018	–0.112	0.132	0.077
Globality	0.243	0.225	0.092	0.189	0.274	0.212
Failure						
Locus	–0.027	–0.120	0.058	–0.075	0.000	–0.007
Stability	0.173	0.119	**0.328***	0.107	0.074	–0.087
Controllability	–0.063	–0.240	–0.100	–0.052	–0.189	–0.124
Globality	0.003	0.102	0.048	–0.095	0.184	0.080

*Spearman's Rho, *: $p \leq 0.05$, **: $p \leq 0.01$*

Table 3 Correlations between attributional dimensions and UEQ scales for situations of success and failure

Group comparison in situations of success and failure

To calculate group comparisons (t tests) two groups were formed with a split at the mean of each attributional dimension.

In *situations of success*, results show significant differences regarding the attributional dimensions of *Locus* and *Globality* (Table 5). Concerning the *Locus* dimension, users with external attribution rated the tested systems significantly more positive than users with an internal attribution on the UEQ scales *Efficiency*, *Stimulation*, and *Novelty*. Participants with high values on the *Globality* dimension rated the tested systems significantly more positive than users with low values on the UEQ scales *Stimulation* and *Attractiveness*.

In *situations of failure*, results show significant differences regarding the dimensions of *Stability* and *Globality* (Table 5). Concerning the *Stability* dimension, users with high expression rated the tested systems significantly more positive regarding *Dependability*. Participants with high values on the *Globality* dimension rated the tested systems significantly more positively than users with low values regarding *Perspicuity*.

Relations between Attribution Styles and UEQ Scales in situations of success and failure

Kruskal-Wallis tests were calculated to investigate whether Attribution Styles are related to UEQ scales.

Regarding *attribution of success* (Table 4), significant differences concerning the UEQ scales *Attractiveness*, *Perspicuity*, *Stimulation*, and *Novelty* were found.

In *situations of failure*, Kruskal-Wallis tests showed no significant differences in any of the six UEQ scales.

Post-hoc tests were calculated to identify the relations between the individual Attribution Styles and UEQ scales (Table 6).

In *situations of success*, the analysis showed significant differences between Attribution Styles and UEQ scales except for *Dependability* and *Efficiency*. Likewise, the Kruskal-Wallis test showed no significant differences for these two scales (Table 4).

Figure 4 shows the mean values of the UEQ scales separately for each Attribution Style. In comparison to the other styles, participants with a *Confident* Attribution Style rated the goal-directed quality aspects *Perspicuity* (M=1.315, SD=1.174), *Efficiency* (M=1.037, SD=0.811), and *Dependability* (M=0.991, SD=0.980) the most positive. Participants with a *Humble* style evaluate *Attractiveness* (M=1.528, SD=0.510) as well as the hedonic quality aspects *Stimulation* (M=1.500, SD=0.632) and *Novelty* (M=1.458, SD=0.828) the most positive. Participants with a *Realistic* style rated the system the most negatively across all UEQ scales.

	Chi^2	df	p
Attractiveness	7.077	2	**0.029***
Perspicuity	6.862	2	**0.032***
Dependability	2.883	2	0.237
Efficiency	4.923	2	0.085
Stimulation	6.858	2	**0.032***
Novelty	8.504	2	**0.014***

Table 4 Relations between Attribution Styles and UEQ Scales in situations of success – results of Kruskal-Wallis test

Success	Expression	n	UEQ Attractiveness	UEQ Perspicuity	UEQ Dependability	UEQ Efficiency	UEQ Stimulation	UEQ Novelty
Locus	low	24	0.602	0.542	0.760	**0.500**	**0.208**	**–0.333**$^{+}$
	high	27	1.178	1.269	0.846	**1.093**	**1.037**	**0.861**
Stability	low	16	0.881	0.625	0.724	0.656	0.547	0.313
	high	35	0.919	1.064	0.843	0.886	0.693	0.293
Controllability	low	32	0.863	0.820	0.844	0.930	0.516	0.195
	high	19	0.982	1.105	0.741	0.618	0.868	0.474
Globality	low	23	**0.522**	0.641	0.743	0.609	**0.217**	–0.022
	high	28	**1.224**	1.161	0.857	0.982	**1.000**	0.563
Failure								
Locus	low	19	1.111	1.250	0.991	1.039	0.855	0.395
	high	32	0.786	0.734	0.696	0.680	0.523	0.242
Stability	low	16	0.751	0.653	**0.417**	0.667	0.611	0.458
	high	35	0.992	1.076	**1.018**	0.894	0.667	0.212
Controllability	low	32	0.910	1.188	0.802	0.823	0.802	0.385
	high	19	0.905	0.694	0.809	0.806	0.509	0.222
Globality	low	23	0.844	**0.500**	0.802	0.844	0.438	0.167
	high	28	0.963	**1.306**	0.809	0.787	0.833	0.417

Note: Bold font means significant difference with $p < 0.05$, $^{+}p < 0.001$

Table 5 Group comparison - high vs. low expression of the attributional dimensions in situations of success and failure

Success	UEQ Attractiveness	UEQ Perspicuity	UEQ Dependability	UEQ Efficiency	UEQ Stimulation	UEQ Novelty
Realist	**0.324**[a,b]	**0.292**[c]	0.532	0.472[(d)]	**0.069**[e,f]	**–0.111**[g]
Confident	**1.158**[a]	**1.315**[c]	0.991	1.037[(d)]	**0.843**[e]	**0.315**[h]
Humble	**1.525**[b]	1.083	0.792	0.833	**1.500**[f]	**1.458**[g,h]
Failure						
Realist	0.633	0.383[(i)]	0.600	0.717	0.250	0.083
Resigned	0.886	1.011	0.807	0.739	0.750	0.375
Confident	1.234	1.375[(i)]	1.024	1.036	0.911	0.411

Note: Superscript letters show significant differences with p < 0.05, bracketed letters indicate previous Kruskal-Wallis test was not significant.

Table 6 Relations between Attribution Styles and UEQ Scales – Post-hoc test

Regarding differences between Attribution Styles and UEQ scales, most significant differences were found between participants with a *Realistic* and a *Confident* Attribution Style, in particular regarding the UEQ scales *Attractiveness* (p=0.025; M=0.324 vs. M=1.158), *Perspicuity* (p=0.012; M=0.292 vs. M=1.315), and *Stimulation* (p=0.044; M=0.069 vs. M=0.843). Thus, participants with a *Confident* style perceive the system as significantly more attractive, easy to understand, and more innovative than persons with a *Realistic* Attribution Style.

Furthermore, there are several notable differences between the *Realistic* and the *Humble* style. These two styles significantly differ regarding the evaluation of *Attractiveness* (p=0.036) as well as *Stimulation* (p=0.017), and *Novelty* (p=0.004). Participants with a *Humble* style perceive the system as significantly more attractive, interesting and innovative than persons with a *Realistic* Attribution Style.

Participants with *Humble* and *Confident* styles differ significantly regarding the hedonic quality aspect *Novelty* (p=0.027; M=1.458 vs. M=0.315). *Humble* users perceive the tested system as more innovative than participants with a *Confident* style.

In *situations of failure*, neither Kruskal-Wallis tests nor Post-hoc tests (Table 6) show significant differences regarding system evaluation and Attribution Styles. Figure 5 shows the mean values of the UEQ scales separately for each Attribution Style.

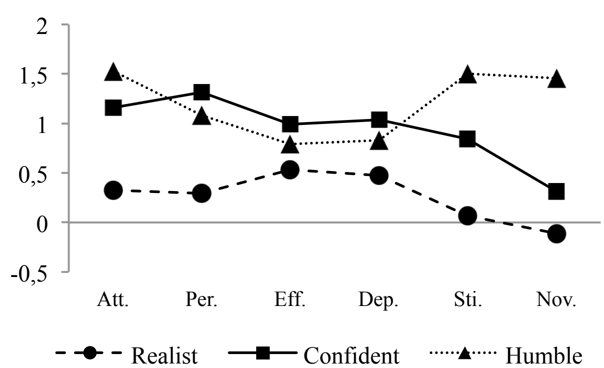

Figure 4 Mean values of UEQ scales for each Attribution Style in situations of success

DISCUSSION

In this study, we investigated whether individual attribution patterns influence users' evaluation of computer systems in usability tests. Our results suggest that this is the case: Particularly regarding situations of success, correlation analysis, group comparisons of attributional dimensions as well as comparisons of different Attribution Styles showed significant differences regarding system evaluations in usability tests. For situations of failure, however, fewer differences were identified.

Influence of Attributional Dimensions

Impact of the Locus dimension on system evaluations (H1)

We hypothesized that system quality aspects are assessed differently depending on whether users show an internal or an external attribution pattern (H1). Looking at the results in Table 5, this hypothesis is mainly confirmed for situations of success: Persons with external attribution patterns rated *Efficiency*, as well as the hedonic quality aspects *Stimulation* and *Novelty,* significantly more positive than persons with internal attribution patterns. Correlation analyses also showed positive significant correlations for *Perspicuity* and *Novelty* (Table 3). Thus, persons who see mainly external reasons for success consistently give better system ratings – explaining their success with good system qualities.

However, for situations of failure this hypothesis was not confirmed by our data. No significant differences regarding *Locus* were observed in any of our tests.

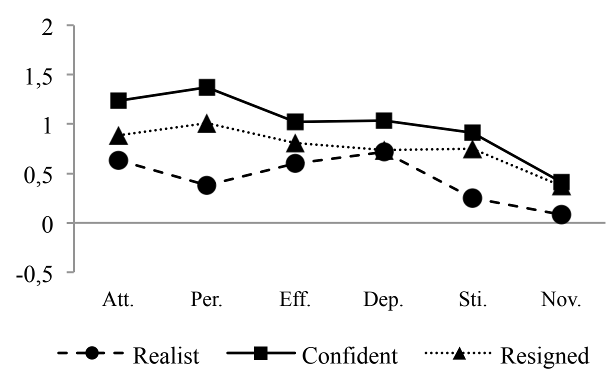

Figure 5 Mean values of UEQ scales for each Attribution Style in situations of failure

Impact of the Controllability dimension on system evaluations (H2)

Furthermore, we hypothesized that persons with low control perception rate a system more negatively than persons with high control perception (H2). This hypothesis could not be confirmed by our data: No significant differences between the *Controllability* dimension and UEQ aspects were identified for success or failure.

Impact of the Stability and Globality dimensions on system evaluations (H3)

Moreover, we hypothesized that in situations of failure persons with high values on the *Stability* and *Globality* dimensions would evaluate systems more negatively (H3). This hypothesis was not confirmed. Quite interestingly, the opposite is the case in situations of failure: Participants with high *Stability* or *Globality* values rated the tested systems significantly more positive regarding *Dependability* and *Perspicuity*, respectively.

Thus, perceiving stable causes for computer problems seems to increase the perceived reliability of a system, even if the use situation as such provides a negative experience. Likewise, global attribution patterns seem to go hand in hand with an increased perception of transparency regarding system design: Perceiving global mechanisms that apply to a wide range of situations and systems seems to increase a sense of understanding. This is a very interesting finding.

Influence of Attribution Styles (H4)

Finally, we hypothesized that individual Attribution Styles have an impact on system evaluation (H4): In particular, we assumed that persons with more favorable Attribution Styles (e.g. *Confident*) would evaluate systems differently than persons with unfavorable styles (e.g. *Resigned* or *Humble*).

In *situations of success*, our results show that indeed persons with a *Confident* style evaluate systems more positively than persons with other Attribution Styles (Figure 4). However, this is true only regarding pragmatic (goal-directed) qualities. Regarding hedonic qualities (*Attractiveness, Stimulation* and *Novelty)*, persons with a *Humble* style gave significantly more positive ratings than persons with other Attribution Styles. This might be due to a greater sense of uncertainty by *Humble* users: As they possibly have a less articulate view when dealing with computers, they therefore might appear more "novel" and interesting to them. Also, being successful in a difficult situation (as computer use is typically challenging for *Humble* users) seems to induce a positive system rating.

Interestingly, persons with *Realistic* styles seem to generally rate computer systems less positively than persons with other styles, regardless of success or failure situations (Figure 4 and Figure 5). However, significant differences were revealed only in situations of success (Table 6).

Regarding situations of failure, the data reveals some obvious differences, which are in line with our theoretical reasoning, albeit not statistically significant. In particular, we suggested that persons with a *Confident* style would evaluate a system more positively, because they blame themselves for the failure and not the system. Looking at Figure 5, this is indeed the case, even though the differences are small.

Limitations

An interesting finding of our study is that for situations of failure, only few significant differences regarding attributions and system evaluation were found. It is hard to argue why attributions might influence system evaluation in success situations, but not in failure situations. A possible explanation is that the systems we tested received a rather positive overall evaluation and failures were seldom or not severe, respectively. Therefore, for future investigations we will consider a study design that brings forth a more balanced experience of success and failure situations (e.g. by giving users easy as well as hard or even unsolvable tasks).

Another limitation is the low variance concerning the attributional dimensions. Especially in situations of failure, Attribution Styles are not so clearly pronounced as in prior studies [15, 22]. Unlike our participants, persons with a *Resigned* style usually experience less control. Perhaps this is due to the homogeneity of the sample with respect to education, age, gender and computer skills, since there is evidence that socio-demographic factors have an impact on attribution processes [21].

Thus, the high level of computer experience among the participants might explain the contradictory results regarding the stability/globality dimensions, as users with high computer expertise understand quite well how computer systems work and therefore interpret failures as consequences of recurring bugs that make the system fail consistently. To investigate the possible impact of these sample characteristics, we re-ran our analyses with computer experience as a covariate. However, this analysis showed no differences, as the variance of the values turned out to be too low. Therefore it is necessary to pay attention to a more heterogeneous sample composition in future studies.

Furthermore, even though conducting usability tests with more than 50 persons already constitutes a rather elaborate study design, the sample size is nevertheless not big, especially regarding group comparisons. Therefore, we will continue to include attribution measures in usability tests to build a larger empirical dataset.

IMPLICATIONS AND FUTURE WORK

At large, our results suggest that there are notable influences of users' attribution patterns on their evaluation of system quality. This is true for both pragmatic and hedonic system qualities: According to our data, 'hard' usability measures such as Perspicuity and Dependability reflect not only stable system characteristics, but also users' traits. Likewise – and maybe less surprising – user characteristics seem to influence 'softer' measures of user experience such as Attractiveness and Stimulation. Especially Locus, Stability and Globality seem to have an impact on how transparent, reliable and generally satisfactory users experience their system interactions.

In our view, this is a very important finding: After all, system designers and HCI experts need to be aware that results of user evaluations might – at least to some extent – reflect their users' general, personal style in experiencing and interpreting situations of computer use rather than the specific qualities of a certain system. Furthermore, if a system is tested by users with specific attribution styles, users with other styles might not share their assessments. Particularly in usability tests with a small sample size (e.g. as recommended by Nielsen [23]) or a very homogenous sample, participants might have similar attribution patterns, which in turn may affect the outcomes of the usability test. For example, the number of bugs reported in usability tests could be related to specific attribution patterns: If a system is tested by users with e.g. mainly external attribution patterns, less bugs might be reported.

Therefore, we recommend including attribution measures in usability tests in research and practice. This can be easily done by integrating a short attribution questionnaire like the one we used into any kind of usability evaluation: I.e., participants are simply asked to fill out the attribution questionnaire in addition to other evaluation measures. That way, researchers and practitioners would be able to include attribution styles as a crucial variable in usability testing and analyze possible effects.

For example, we are currently preparing to integrate the attribution questionnaire into the survey system of a well-known travelling portal. Approximately 600 of the portal members regularly participate in design and usability surveys. Analyzing the usability data with respect to attributions will broaden our view on the influence of attribution patterns on system evaluations.

Additionally, we are planning another laboratory study with a more heterogeneous sample composition, especially concerning computer experience and skills. Furthermore, we will use test applications that differ strongly in terms of the pragmatic quality aspects *Perspicuity*, *Efficiency,* and *Dependability* to investigate their relation with attribution styles.

Moreover, it's an important question for future research whether certain *design decisions* in the design of interactive systems are perceived and evaluated differently by users with different attribution patterns. E.g., a high degree of user freedom vs. a strongly guided user interface might be perceived differently by users with internal vs. external locus of control.

The present study cannot answer this question, as we deliberately used a wide range of applications in our usability tests to analyze whether there are any correlations between attributions and system evaluations at all. However, it will be a logical next step in HCI-related attribution research to identify design patterns that are specifically supportive and provide a good user experience for users with certain attribution styles.

To sum up, our study can be seen as a first investigation of the relations between attribution patterns and system evaluation. Attribution research is a novel field in HCI; especially the influence of attribution styles on system evaluation has not been researched yet. Our results hint that there is an association indeed, which has to be clarified and investigated in more detail. Nevertheless, given the current state of research, we believe that it is important for researchers and practitioners to bear in mind that system evaluations might have been influenced by attribution patterns when conducting and interpreting usability tests.

ACKNOWLEDGMENTS

We thank our anonymous reviewers for their helpful comments. Special thanks go to Reviewer 3 for his/her suggestions how to refine our statistical analyses.

REFERENCES

1. Janos Abonyi and Balázs Feil. 2007. *Cluster Analysis for Data Mining and System Identification*. Birkhäuser Boston.
2. Lyn Y. Abramson, Martin E. Seligman, and John D. Teasdale. 1978. Learned helplessness in humans: critique and reformulation. *Journal of abnormal psychology* 87: 49–74. http://doi.org/10.1037/0021-843X.87.1.49
3. Lyn Y. Abramson, Gerald I. Metalsky, and Lauren B. Alloy. 1989. Hopelessness depression: A theory-based subtype of depression. *Psychological Review 96*, 358–372. http://doi.org/10.1037/0033-295X.96.2.358
4. Johann Bacher, Andreas Pöge, and Knut Wenzig. 2010. *Clusteranalyse: Anwendungsorientierte Einführung in Klassifikationsverfahren*. Oldenbourg Wissenschaftsverlag.
5. Jacob Cohen. 1988. *Statistical power analysis for the behavioral sciences*. Erlbaum, Hillsdale. http://doi.org/10.1234/12345678
6. Oliver Dickhäuser and Joachim Stiensmeier-Pelster. 2002. Erlernte Hilflosigkeit am Computer? Geschlechtsunterschiede in computerspezifischen Attributionen [Learned helplessness in working with computers? Gender differences in computer-related attributions]. *Psychologie in Erziehung und Unterricht* 49: 44–55.
7. Oliver Dickhäuser and Joachim Stiensmeier-Pelster. 2002. Entwicklung eines Fragebogens zur Erfassung

computerspezifischer Attributionen. *Diagnostica* 46, 2: 103–111. http://doi.org/10.1026//0012-1924.46.2.103

8. Carol S. Dweck. 1999. *Self-theories: Their role in motivation, personality, and development.* Psychology Press, Hove. http://doi.org/10.1007/BF01544611
9. Valerie S. Folkes. 1988. Recent Attribution Research in Consumer Behavior: A Review and New Directions. *Journal of Consumer Research 14*, 548. http://doi.org/10.1086/209135
10. Friedrich Försterling. 2001. *Attribution. An introduction to theories, research, and applications*. Psychology Press, Hove.
11. Nicolas Gillet, Robert J. Vallerand, and Elisabeth Rosnet. 2009. Motivational clusters and performance in a real-life setting. *Motivation and Emotion* 33, 1: 49–62.
12. Sascha R. Guczka and Monique Janneck. 2012. Erfassung von Attributionsstilen in der MCI – eine empirische Annäherung. *Reiterer, H., Deussen, O. (Hrsg.), Mensch & Computer 2012: interaktiv informiert – allgegenwärtig und allumfassend!?* In Reiterer, H., Deussen, O. (Hrsg.), 223–232.
13. Stephanie J. Hanrahan, J. Robert Grove, and John A. Hattie. 1989. Development of a questionnaire measure of sportrelated attributional style. *International Journal of Sport Psychology* Vol 20(2): 114–134.
14. Andreas Hinderks, Martin Schrepp, and Jörg Thomaschewski. Analyzing the User Experience Questionnaire. Retrieved 22 September 2015 from http://www.ueq-online.org/
15. Monique Janneck and Sascha R Guczka. 2013. The Resigned, the Confident, and the Humble: A Typology of Computer-Related Attribution Styles*,* In *Human Factors in Computing and Informatics (Proc. SouthCHI 2013)*.
16. Robin H. Kay. 1990. The relation between locus of control and computer literacy. *Journal of Research on Computing in Education* 22, 2: 464. Retrieved from https://vpn.cut.ac.cy/
17. Grace Kovenklioglu and Jeffrey H. Greenhaus. 1978. Causal attributions, expectations, and task performance. *Journal of Applied Psychology 63*, 698–705. http://doi.org/10.1037/0021-9010.63.6.698
18. Bettina Laugwitz, Theo Held, and Martin Schrepp. 2008. Construction and Evaluation of a User Experience Questionnaire. In *HCI and Usability for education and work, proceedings*, 63–76.
19. Robert G. Lord and Jonathan E. Smith. 1983. Theoretical, Information Processing, and Situational Factors Affecting Attribution Theory Models of Organizational Behavior. *Academy of Management Review* 8, 1: 50–60.
20. Herbert W. Marsh, Barbara M. Byrne, and Richard J. Shavelson. 1988. A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. *Journal of Educational Psychology* 80: 366–380. http://doi.org/10.1037/0022-0663.80.3.366
21. Adelka Niels, Sascha R. Guczka, and Monique Janneck. 2015. Computer-related Causal Attributions: The Role of Sociodemographic Factors. In *Proceedings of 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, AHFE 2015*. Elsevier, pp. 2483–2490.
22. Adelka Niels and Monique Janneck. 2015. *Computer-Related Attribution Styles: Typology and Data Collection Methods.* In J. Abascal et al. (Eds.), *INTERACT 2015, Part II, LNCS 9297*, pp. 274–291.
23. Jakob Nielsen and Thomas K. Landauer. 1993. *A mathematical model of the finding of usability problems*. In CHI '93 Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems, pp. 206-213.
24. Christopher Peterson, Amy Semmel, Carl von Baeyer, Lyn Y. Abramson, Gerald I. Metalsky, and Martin E. Seligman. 1982. The attributional Style Questionnaire. *Cognitive Therapy and Research* 6, 3: 287–299. http://doi.org/10.1007/BF01173577
25. Martin Schrepp, Andreas Hinderks, and Jörg Thomaschewski. 214AD. Applying the User Experience Questionnaire (UEQ) in Different Evaluation Scenarios. In *Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience*, Marcus Aaron (ed.). Springer International Publishing, Springer International Publishing, 383–392.
26. Ralf Schwarzer. 1994. Causal attributions as health-related cognitions. In *F. Försterling & J. Stiensmeier-Pelster, Attributionstheorie. Grundlagen und Anwendungen*, pp. 213-234.
27. Astrid M. Sølvberg. 2002. Gender differences in computer-related control beliefs and home computer use. *Scandinavian Journal of Educational Research* 46: 409–426.
28. Bernard Weiner. 1985. An attributional theory of achievement motivation and emotion. *Psychological review* 92, 4: 548–573. http://doi.org/10.1037/0033-295X.92.4.548
29. Bernard Weiner. 1974. Achievement motivation and attribution theory. *General Learning Press, Moristown*.