Safety and Health at Work through Persuasive Assistance Systems

Matthias Hartwig and Armin Windel

Federal Institute for Occupational Safety and Health - Division
"Products and Work Systems", Friedrich Henkel 1-25 D-44149 Dortmund, Germany
{hartwig.matthias,windel.armin}@baua.bund.de

Abstract. In working environments, violations against safety regulations like the use of personal protective equipment pose a significant threat to well-being and health of working people. A laboratory study investigated the potential of different computer generated feedback forms encouraging users to wear protective equipment (PPE), even when this PPE hinders their primary work task, thus threatening their financial compensation. The results show a substantial increase in usage of PPE when being confronted with a persuasive designed feedback like a traffic light or an emotional expression of a virtual avatar. In contrast, solely informative feedback showed no significant impact of safety behavior. In summary, the findings indicate a potential of persuasive technology for occupational safety affairs and underline the importance of the outward appearance of computer generated persuasive messages.

Keywords: Adaptive Work Assistance Systems, Chances and Risks, Evaluative Feedback, Persuasion, Safety and Health.

1 Theory

By now, there is extensive knowledge about safe and healthy behavior at work: Education, information, rules and regulations are used to ensure this behavior in operational practice, for example the use of personal protective equipment (PPE). The effectiveness of these measures varies substantially, depending on the field of application. In sum however safe and healthy behavior is often neglected in practice. This suggests a substantial deficit in the perceived necessity of such behavior, in spite of all existing threats of penalty.

On organizational level, conscious infringements against safety regulations are often driven by economical motives. On individual level, reasons are divergent. Reason [1] classifies "unsafe acts" into unintentional errors like slips and mistakes, and intentional violations, which are based on a conscious decision to act against the regulation. These conscious violations often pose an especially high risk, because they commonly form a habit and will most likely be repeated in every similar situation. As a result, the individual risks for an accident to happen add up over time. Motives for these intended violations can be twofold: Firstly, the person could consider the safety behavior to be inappropriate or useless and has no intention to perform it, which

V.G. Duffy (Ed.): DHM/HCII 2013, Part II, LNCS 8026, pp. 40–49, 2013.

means the person has a negative attitude towards the respective behavior. Secondly, the person could consider the safety behavior per se to be useful, but its negative aspects outweigh the stimulus for acting or are more salient in the specific situation. These aspects can range from additional effort, slower working progress to less thrilling experiences at work. In this case, the person has a positive attitude towards the behavior itself, but in the specific situation, it is not strong or salient enough to activate the associated behavior. The second constellation is very relevant in every day life, because the prevention of accidents is rarely perceived as the primary goal of work, and so the risk and the subjective relevance of safety behavior are pushed to the background in the mind of the working person. The probability for violations is therefore especially high, when safety behavior is perceived as hindering for the working goals. To counteract these tendencies, a solution might be to remind the user of the relevant safety behavior and encourage it, preferably in the very moment the behavior is indicated.

Modern man-machine systems have the potential to provide such assistance, particularly those oriented to the technology vision of ambient intelligence. This technology paradigm is based on the idea of "ubiquitous computing" by Marc Weiser [2] and is characterized by Aarts [3] by the central features context awareness, personalization, adaptive behavior) and anticipation. In the working environment, these are called adaptive work assisting systems (AWAS,[4]) Applied to the described scenario, these properties enable the system to reduce violations by (1) being aware of the behavior of the user, (2) evaluating it autonomously regarding violations and (3) presenting evaluative feedback that changes the users behavior. Computer interfaces, purposely designed to change the behavior of the user, can be subsumed under the term persuasive technology [5]. While there are numerous approaches in field of persuasive technologies to investigate the applications for e-commerce, environmental protection or private healthcare, there are only few efforts made on how persuasive technology can be applied in the working environment. Furthermore, it is still unclear whether the outward appearance of feedbacks is critical for the persuasion. Two approaches are particularly relevant for the described scenario: (a) forms that are already associated with action stimuli from everyday life such as traffic lights, and (b) anthropomorphic interfaces such as animated virtual agents. Reeves and Nash [6] were able to show that users involuntarily attribute human characteristics to computer interfaces with human-like appearances. Therefore, computers can provide similar social cues as human do. The question, if this implies similar effects and action mechanisms as in social persuasion, is subject of an ongoing debate (for examples see [7, 8]).

The German Federal Institute for Occupational Safety and Health (BAuA) therefore conducted a laboratory experiment to investigate the potential of those different persuasive feedback forms to facilitate the use of personal protective equipment. The study is part of its current research focal point Ambient Intelligence (AmI), evaluating chances and risks of new adaptive technologies in the working environment. Testing the effects of persuasive feedback for safety and health behavior requires a setting that meets certain requirements. Participants should be able to accomplish the task without special knowledge, the need for PPE should be easily comprehensible without exposing participants to real hazards and the setting should be static, so the feedback on a monitor can always be seen by the participants. Taking these aspects into account, a simulation of a simple electrical engineering task was chosen as working task, while usage of isolating gloves was selected as corresponding safety behavior.

2 Procedure and Materials

All participants were given detailed standardized instructions on their task to manually build ten electronic circuits correctly and as quickly as possible according to a step by step guide on the monitor. They were also informed that during certain working steps there is a risk of an electric shock (which was in fact not the case). The subjects were instructed to wear insulated gloves as PPE in these operations. Usage of these gloves was the primary dependent variable of the experiment. The thick and stiff work gloves impaired and slowed down the filigree task of building the circuits significantly, creating a conflict between the two objectives given.

This conflict was exacerbated by two means. First, a bonus of 5 Euro (~ $6) for task completion in less than the average time was promised. This average time was not specified to the subjects; in fact, all participants received this “bonus”. In contrast, use of gloves was not explicitly taken into account for payment. Secondly, after 30% of the circuits have been built, all subjects received a computer generated message, stating that they are about 2 Minutes slower than the average so far (regardless of their actual speed) and that their current working speed would therefore not be sufficient to receive the bonus.

Fig. 1. Two different negative (left) and positive (right) Avatar feedbacks

In the control group, the subjects worked on these tasks without additional information on their PPE use. In the three other experimental groups, different forms of feedback on their use of gloves appeared on one half of the instruction monitor at each corresponding step. All feedbacks were accompanied by a very short bell sound, to make sure that they were recognized. In the experimental group "text" a short, purely informative held writing appeared, either "gloves used" as positive feedback, or "Please wear gloves" as negative feedback. In the experimental group "traffic light" the same text was presented, accompanied by a picture of a traffic light, displaying either green (positive) or red (negative) light. Finally, in the experimental group "virtual agent" an anthropomorphic virtual agent was shown, which presented either one of two positive predefined expressions (joy) or one of two negative (anger or sadness, all four different expressions are shown in Figure 1). To make the agent-based feedback more vivid and natural, in this experimental group the accompanying text was not held constant, but varied randomly between ten positive - affirmative (e.g. "very good, Gloves used") or negative - prompting messages (“forgetting something?”). The study was designed as a between-subject experiment; each subject took part in only one of these experimental conditions. The feedback was realized by the experimenter using the Wizard of Oz- technique. During the study, the experimenter - invisible to the subjects – recorded the “violations”, which is the number of working steps were gloves were not use although the subject was supposed to. Other dependent variables were speed and quality of task performance, measured by the time the participants needed to complete the circuits and the number of correctly reconstructed circuit boards. As part of a manipulation check, a questionnaire asked (1) for the general perceived time pressure, as well as (2) how hindering the gloves were. Other questions asked (3) as how necessary the gloves were perceived, (4) how helpful the feedback was during the task and (5) whether the feedback had an impact on the participants’ decision, to wear or not to wear the gloves. These items were used to understand the psychological mechanisms of action of the different feedback forms. Psychological acceptance was measured by the items (6) “I would not comply with such a feedback on principle”, (7) “I could imagine an automated feedback system on daily work” and (8) “I could imagine an automated feedback system during training time”. All Items were answered based on 5 point Likert scale (ranging from 1 = “does not apply at all” to 5 = “applies completely”). Participants answered the Neo-FFI [9] in its German translation [10], a personality questionnaire on the big five to rule out the influence of personality traits on experimental behavior.

2.1 Participants

The total sample of external subjects consisted of eighty persons. Five participants were ruled out from the statistical analyses. Two did not meet the internal time limit of the experiment, one was unable to use the gloves due to a bandaged hand and two persons were no native speakers and had an insufficient language ability to understand the instructions on the screen. Resulting seventy-five subjects (38 females, 37 males; age 25,75; *SD* 4,26) in the final sample were randomly assigned to one of four experimental conditions (feedback in form of an virtual avatar (n= 19), a traffic light (n= 18), a text message (n=18) or no feedback at all (n= 20). All subjects were mainly recruited on the internet and through announcements at nearby universities. All resulting participants were native German speakers. The experiment lasted between 120 and 150 Minutes, for which the participants were paid 25 € (approximately $ 32 US at the time the study was conducted).

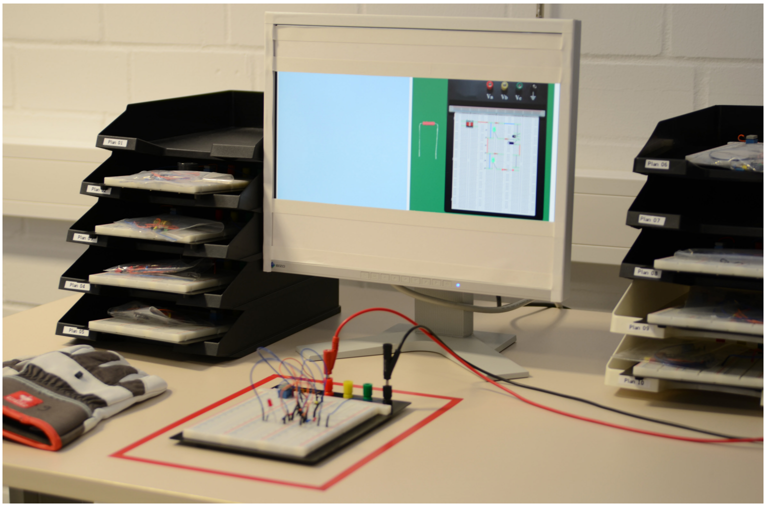

Fig. 2. Experimental work station with gloves (left), circuit board (bottom) and instruction monitor (center)

2.2 Hypotheses

Based on the theoretical considerations, it was hypothesized that the feedback in general has an impact on safety behavior, resulting in the experimental groups differing significantly in terms of glove use (hypothesis 1). More specifically, it was assumed that the informative text-based feedback increases the salience of safety behavior, thus increasing the probability of occurrence (hypothesis 2a). The same effect should be achieved by the persuasive designed feedback forms so we assumed that the Traffic Light and Virtual Avatar – groups perform fewer violations than the control group (hypothesis 2b). In addition, it was hypothesized that the effect of persuasive feedback was greater than the one of a purely informative feedback. This should result in more violations in the text group than in traffic light and virtual avatar group (hypothesis 3). Finally, it was assumed that the effects of the two persuasive feedbacks to be based on different psychological mechanisms, reflected in different impacts (hypothesis 4). The questionnaire was designed in order to give information on these different mechanisms.

3 Results

3.1 Descriptive Statistics and Tests on Normal Distribution

The safety behavior was operationalized by counting the number of operations where the participant was not wearing gloves although supposed to (called "violation"). Before the message, out of 11 possible violations, participants score an average of 2,19 (*SD* 3,486, Range 0-11). After the message, the average number of performed violations was 6,43 (*SD* 9,54, Range 0-23) out of 23 possible violations. 41

participants scored zero violations, always wearing the gloves during the corresponding steps. A Kolmogorov-Smirnov test therefore showed an in-normal distribution of violations (*p*. = 0.000), skewed to the right.

Concerning working speed, it took the participants on average 3509 seconds (SD 747 sec. Range 1928-5550 sec.) to complete all ten circuits. After the manipulated working speed report between boards three and four, average time to complete the remaining seven boards was 2182 sec. (*SD* 537 sec., Range 1219-3697 sec.). Both curves were normally distributed (KS-test, $p > .05$). The quality of work was measured by counting the correct assembled and working circuits for each participant. The average number of working circuits was 6,36 (*SD* 1,73, Range 2-10).

In the questionnaire, participants reported a mediocre time pressure (mean 3,68, on a scale between 1-5, *SD* 0,774). As intended, the gloves were stated as very hindering for the task at hand (mean 4,41, *SD* 1,25) and not necessary (1,96, *SD* 1,27). In all feedback-groups, it was reported as being mediocre useful (3,00 *SD* 1,09) and disturbing (2,80 *SD* 1,41), as well as having a mediocre effect on the participants behavior (3,13 *SD* 1,39. Conscious reactance (Item 6, "I would not comply with such a feedback on principle") was stated relatively low with a mean of 1,75 (*SD* 9,89), while acceptance for use in daily work/training time was reported relatively high (mean of 3,87 *SD* 0,89 and 4,24 *SD* 0,816 respectively).

3.2 Manipulation Check: Conflict

First, it was checked whether the intended conflict between working speed and safety is reflected in the data. To test this, working speed and violations were correlated. Kendall's-Tau coefficient was used the because of the in-normal distribution of violations. Confirming our intentions, the test showed a significant negative correlation (-.244, $p. < .05$) between total working time and violations, which means using gloves led to increased working time.

3.3 Performance and Safety Behavior

The **Hypothesis 1** states that text-based, purely informative feedback would lead to fewer violations than in the control group without any feedback. To test this, violations before and after the manipulated working speed message were submitted to a Univariate Analyses of Variance (ANOVA) each, with Feedback Form as independent variable. ANOVA was chosen due to its error-robustness, though the in-normal distribution of violations infringes its mathematical preconditions. The Welch-Value was calculated additionally in order to rule out distortion caused by deviating variances. and ensure the results.

After the message, the experimental groups differed significantly from each other regarding glove usage. (ANOVA: $F = 3{,}445$, $p = 0{,}021$). The additionally conducted Welch Test revealed a significant difference as well ($F = 3{,}286$, $p = 0{,}031$). The average score of violations in the experimental groups are shown in Figure 3.

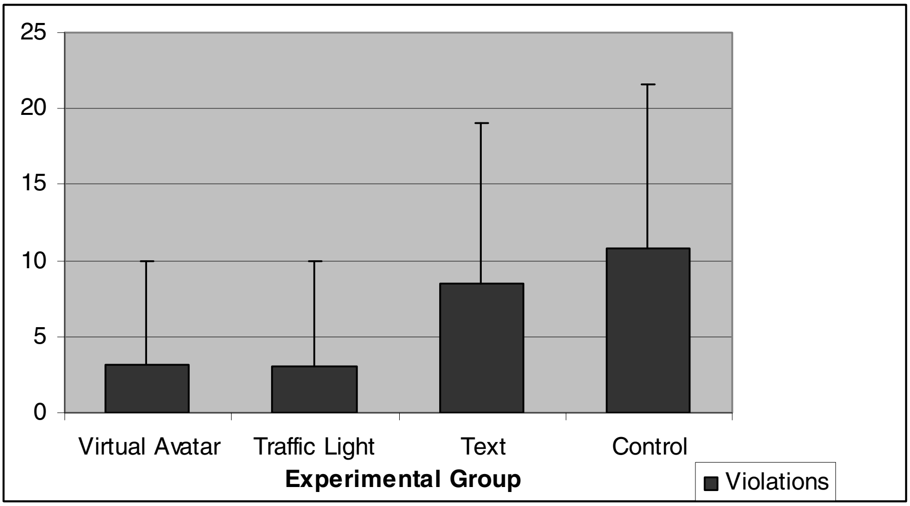

Fig. 3. Average violations and standard deviations for the experimental groups

On account of the floor effect concerning the distribution of violations, the *F* values are comparatively low for the given percental discrepancy of incidents (e.g. in the control group, roughly three times more violations occurred compared to the Virtual Agent group). No post-hoc pair comparisons show significant results. After balancing out contentual relevance and formal statistical aspects, it was decided to calculate the paired comparisons between the experimental groups with t-Tests. In order to avoid a summation of alpha-errors, we tested only those pairs which are part of the hypotheses.

Hypothesis 2a states that fewer violations occur in the textual feedback group compared to the control group. A one tailed t-test results in a $p = 0{,}516$ ($t = -0{,}656$), so the hypothesis is not confirmed.

Hypothesis 2b postulates that fewer violations occur in the persuasive feedback groups compared to the control group. A one tailed t-test reveals a significant difference ($t = 2{,}252$, $p = 0{,}002$) between the averages, confirming the hypothesis.

In **Hypothesis 3**, it was assumed that the persuasive feedback groups have fewer violations compared to the textual feedback group. The respective one tailed t-test shows a significant difference as well ($t = 2{,}252$, $p = 0{,}028$), so hypothesis 3 is confirmed.

Hypothesis 4 assumed the two kinds of persuasive Feedback to have a different impact on violations as well. A two tailed t-test reveals no significant result though ($t = 0{,}093$ $p = 0{,}926$), leaving hypothesis 4 unconfirmed.

To gain insight about the psychological impact forms, the questionnaire ratings were analyzed for the two persuasive designed feedback forms via t-tests. On three items, the two feedback groups showed significant discrepancies. The gloves were rated as significantly more hindering when confronted with feedback by traffic lights ($t = -2{,}444$, $p = 0{,}025$). In the virtual agent group, feedback was assessed as more disturbing (t = 2,081, p = 0,045). Finally, participants rated the traffic light as more acceptable as a feedback system during training time (t = -2,794 p = 0,008). Two more items showed noticeable results: feedback by a virtual avatar was numerically rated having a greater impact on safety behavior, while in the traffic light group, the glove itself were assessed as more necessary. In both items however, the two groups did not differ significantly by a small margin (t = 1,825, p = 0,078 and t = -1,931, p = 0,064, respectively). Concerning work speed and work quality, ANOVAS revealed no significant differences between the experimental groups regarding working time and number of working circuits, neither before nor after the interim message.

4 Discussion

The study presented investigates the influence of different feedback forms on safety violations. A dynamic feedback on safety behavior based on persuasion was expected to reduce the number of safety violations, even when this means acting against ones own financial interests. In a laboratory study, participants performed an electric task, building electronic circuits. They worked under time pressure, being told that fast performance would grant them a financial bonus, and asked to simultaneously use unhandy gloves as personal protective equipment on specific operations, slowing down the working process. While working on the task, they either received no evaluating feedback on their PPE use, a purely informative text message, a picture of a traffic light or the picture of a virtual avatar showing an emotional expression. Safety behavior was measured, counting the numbers of actual PPE-use during relevant operations. Additionally, working speed and quality of work were recorded, while a questionnaire gauged various attitudes concerning task, feedback and safety behavior. Taking quality of work and speed into account, no significant differences between the experimental groups were found, suggesting that such forms of feedback on safety behavior have no substantial impact on the primary task. Concerning safety behavior, displaying feedback per se is not sufficient to significantly reduce safety violations. However persuasive designed feedback, either in form of a traffic light or a virtual avatar showing an emotional expression, had not only a statistically significant but also a substantial impact on safety behavior, reducing the violations occurring in the absence of any feedback by roughly 60 %.

The results of the presented study suggest that purely informative feedback on safety behavior is not sufficient to reliably prevent violations, even when it is presented at the most relevant moment. Instead, the appearance of the feedback seems to be a crucial factor when it comes to impact on user behavior. Intriguingly, in this study the feedback did not only change behavior of the participants, but it successfully persuaded them to act directly against their own financial interests. This strongly suggests that persuasive designed feedback does not only work as reminder that triggers a behavior to which people are motivated anyway, but can act as a factor of its own in decision making. For occupational practice, these results strongly advocate to consider the human psychology when using signs and guidelines to ensure safety behavior. It may not be enough to make sure these signs are readable and understandable. On the contrary, the actual impact on the behavior might depend on its outward appearance, especially when there is no thread of punishment or it is perceived as unlikely to happen. Since the persuasive feedback in this study had an impact on behavior, without being bolstered by the thread of sanctions, one can assume that persuasive technology can be used even where certain safety behaviors are desirable, but prohibitions or sanctions are inappropriate. This was considered to be especially important because presently, there are only few “gentle” methods available to support safety behavior without using prohibitions and sentences. At the same time the results imply that a responsible application of persuasive technology is mandatory. Because of its impact on behavior, this kind of technology works somewhere in between assisting the users and manipulating them. Future research efforts should evaluate this continuum and define the terms of an ethical action guiding versus an unacceptable violation of autonomy.

When comparing the two persuasive designed feedback forms (traffic light and virtual avatar) with each other, no significant difference were found concerning direct impact on safety behavior. The questionnaire however, indicates that there were substantial differences concerning how the feedback influences the user psychologically. Participants rated anthropomorphic feedback significantly more disturbing, but also as having a larger impact on decision making, while the traffic light feedback is rated as more acceptable during vocational adjustment. Finally, presentation of traffic light feedback leads to more negative rating of the safety behavior itself. Taken together, these findings might suggest that anthropomorphic feedback is experienced as more intrusive and draws more attention, while the traffic light triggers a well known association (red – “dangerous”, green – “ok”) without much thought about the traffic light itself.

It remains unclear however, if these differing subjective measures reflect distinct psychological mechanisms of behavior change. If so, the different feedback forms might show distinctive impact on behavior under different circumstances, for instance longer working time, higher cognitive demands of the primary task, or regarding persistence of behavior change. The last aspect is considered of particular importance. All forms of feedback scored a substantially higher acceptance rate for training period than for permanent use, which advocates a temporary use, where achieving a permanent behavior change would be mandatory. One could assume hypothetical explanations for the superiority of both feedback types. On the one hand, the observed better (in this case: less hindering) assessment of the safety behavior when confronted with the virtual avatar could lead to fewer violations, even after the feedback is no longer presented. However, it is not clear how continual this assessment would be. On the other hand, the anthropomorphic feedback, being perceived as more disturbing and having more impact, might represent a consciously perceived external motivation. If so, the increase of safety behavior would be displayed to receive the positive feedback, and would therefore go back when the external reward ceases. It might even have an effect of undermining intrinsic motivation [11]. As a result, a planned follow up study will focus on more accurate insights (1) on the psychological mechanisms how different forms of automatic feedback influence user and (2) the stability of behavior changes.

Acknowledgements. We would like to thank the laboratory unit and especially Ulrich Hold and Nina Schelter for invaluable assistance and support during the realization and implementation of the study.

References

1. Reason, J.: The Human Contribution. Unsafe Acts, Accidents and Heroic Recoveries. Ashgate Publishing (2008)
2. Weiser, M.: The computer for the twenty-first century. Scientific American 265(3), 94–104 (1991)
3. Aarts, E.H.L., Harwig, H., Schuurmans, M.: Ambient Intelligence. In: Denning, J. (ed.) The Invisible Future, pp. 235–250. McGraw Hill, New York (2001)
4. Windel, A., Hartwig, M.: New Forms of Work Assistance by Ambient Intelligence. In: Paternò, F., de Ruyter, B., Markopoulos, P., Santoro, C., van Loenen, E., Luyten, K. (eds.) AmI 2012. LNCS, vol. 7683, pp. 348–355. Springer, Heidelberg (2012)
5. Fogg, B.J.: Persuasive Technology: Using Computers to Change What We Think and Do. Morgan Kaufmann (2003)
6. Reeves, B., Nass, C.: The Media Equation: how people treat computers, television, and new media like real people and places. University Press, Cambridge (1996)
7. Roubroeks, M.A.J., Ham, J., Midden, C.J.H.: When artificial social agents try to persuade people: The role of social agency on the occurrence of psychological reactance. Social Robotics 3(2), 155–165 (2011)
8. Schulman, D., Bickmore, T.: Persuading Users through Counseling Dialogue with a Conversational Agent. In: Proceedings of Persuasive Technology 2009, Claremont (2009)
9. Costa, P.T., McCrae, R.R.: Normal personality assessment in clinical practice: The NEO Personality Inventory. Psychological Assessment 4(1), 5–13 (1992)
10. Borkenau, P., Ostendorf, F.: NEO-Fünf-Faktoren-Inventar (NEO-FFI) nach Costa und McCrae (S. 5-10, 27-28). Hogrefe, Göttingen (1993)
11. Deci, E.L., Koestner, R., Ryan, R.M.: A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation. Psychological Bulletin 147, 627–688 (1999)