Error Training: Replication and the Function of Exploratory Behavior

Tanja Dormann & Michael Frese
University of Giessen

A study of 30 psychology students replicated and extended the findings of Frese et al. (1991) on the superiority of error training, a procedure in which trainees are exposed to many errors. The hypothesis was that error training would lead to exploration and this would increase performance. The error training group performed better than the group that received training in which they were not allowed to make any errors. Students in both the error training and the error avoidant groups performed better if they explored. Exploration was done in contrast to the instruction in the error avoidant group. The data suggest that it is necessary to look into the question of whether errors instigate exploration.

INTRODUCTION

There has been a substantial advance in the literature on training for human-computer interaction since the mid 1980s. Many authors have emphasized the importance of exploration, allowing the use of an active approach to learning (Carroll, Mack, Lewis, Grischkowski, & Robertson, 1985; Frese et al., 1988; Greif & Keller, 1990). In a separate development, error training has been proposed as an effective procedure for learning a computer system (Frese et al., 1991; Greif & Janikowski, 1987; Irmer, Pfeffer, & Frese, 1991). Error training implies that trainees are forced to make many errors and are encouraged to learn from them. Greif and Janikowski (1991) have suggested, at least implicitly, that error training and exploratory training are working via the same mechanisms. They have, therefore, called their procedure “exploratory learning through errors.” Before discussing the relationship between error training and exploratory training, the two forms should be described.

Correspondence and requests for reprints should be sent to Michael Frese, Department of Psychology, University of Giessen, Otto-Behaghel-Strasse, 10F, D-35394 Giessen, Germany.

In error training, the subjects are encouraged and forced to make errors. In one study, there was one period of enforced error training that involved giving the subjects problems that exceeded their level of expertise (Frese et al., 1991). To reduce the amount of frustration that may result from such a procedure, heuristics were introduced. For example, statements such as "There is a way to leave the error situation" or "I have made an error. Great!" were introduced in order to support the perception that errors are necessary to the learning process. Error training proved to be superior to other forms of training such as a sequential training procedure without a chance to make errors (Frese et al., 1991), a training procedure mimicking a tutorial (Greif & Janikowski, 1987), and an experienced trainer in a computer skills training school using his normal procedure (Irmer et al., 1991).

Exploration in training implies that the trainee takes on the role of the learner by trying out different solutions to the problem at hand. However, this is not to be confused with blind trial-and-error, because exploration may be based on hypotheses and guided by some kind of mental model (Carroll, Mack, Lewis, Grischkowksi, & Robertson, 1985; Greif & Keller, 1990). In exploratory training programs, usually some material explaining the basics of the system is provided. In addition, participants are encouraged to browse through the computer system independently and without much outside help, to get familiar with its functioning. This active approach is of particular importance for learning to deal with real-life tasks (Carroll et al., 1985; Frese & Zapf, 1994). Moreover, a person gets to know more system states if he or she explores.

There is some overlap between error training and exploration. Exploratory training implies that one makes errors because the environment is not structured and the information base is minimal. Error training, on the other hand, implies a certain amount of exploration because the trainee has to pursue actively the development of knowledge, and one has to search for one’s own ways of solving problems and reacting to novel and unknown system situations.

Possibly, the performance advantages of error training are due to the exploration that is inherent in such training. An error may actually instigate exploration. This would mean that a trainee in an error training group gets to know more of the system by having to explore it more actively. Another potential mechanism that leads to the superiority of error training is that one learns to deal rationally with the emotions involved in making errors. The heuristics just discussed may reduce frustration due to errors, thus leading to better emotional strategies for dealing with errors in the performance test. A third mechanism might be that one learns to deal with error situations more effectively after error training.

The first of these three potential mechanisms through which error training improves learning was studied in this experiment. The experiment had two goals. One goal was to extend and generalize the experimental evidence of Frese et al. (1991) to a different, more complex task (a statistics program). The other goal was to examine the mechanism of exploration as a potential explanation of the effectiveness of error training.

Similarly to Frese et al. (1991), the hypothesis was that error training would lead to higher after-training performance than training that would not allow error to appear. This evidence should be apparent for complex tasks because more errors can occur when performing such tasks, more system knowledge is needed, and these tasks are not mere replications of what one has learned in training.

Error training is contrasted with error avoidance training, a procedure that allows no exploration, in principle. Error avoidance training mimics tutorials based on Skinner's (1968) programmed learning. This type of training is based on his argument that errors are punishments that need to be avoided, thus error avoidant training should lead to better performance. No errors can appear in such training because the procedures for how to solve the tasks are prescribed in detail; so long as the subjects follow these explicit instructions, they cannot make any errors.

However, as Carroll et al. (1985) and Carroll and Mack (1983) have observed, computer users do not like to comply with such a procedure without exploration. We, therefore, hypothesized that in any error avoidant group, there would be a proportion of subjects who do not comply with the instruction to avoid errors and do not follow the prescribed procedures. Furthermore, we expected those noncompliant subjects of the error avoidant group to be better performers precisely because they explored. This special group’s performance may be similar to that of the error training group.

METHODS

Subjects

Subjects were 30 psychology students, randomly assigned to either the error training or the error avoidant group; 14 were male, 16 were female, and ages ranged from 20 to 36 years (with an average age of about 26 years). Psychology subjects had to be taken in this case because basic knowledge in statistics and the mainframe version of SPSS/X was needed; all subjects had attended courses in statistics and SPSS/X. However, most of them did not work with SPSS/X (26 had never worked with SPSS/X, 2 seldom, 2 sometimes, none often). None had any experience with SPSS/PC. Subjects’ average computer experience was 2-1/2 years, mainly with word processing programs.

Experimental Procedure

Subjects received individual training on the computer. The software to be learned was the statistical package SPSS/PC. The goal of the training was to learn some rudimentary SPSS/PC commands necessary to perform statistical calculations.
The whole experiment lasted for 3 hours per subject, with training time being the same for all subjects. The subjects participated in a 2-hour training program, and after the training they were tested for 1 more hour. The training involved three groups of tasks: creating an input layout for a data-entry procedure, entering the data, and computing a correlation. The subjects were asked to think aloud. According to Ericsson and Simon (1980) this is an appropriate method to gain insight into the learner's thought processes.

The error avoidant group (n = 15) received training designed to reduce the chances of making errors. They were given written instructions specifying each step and command to be used. If they did happen to make any error (e.g., mistyping a command) it was quickly corrected by the experimenter. No further information was given. Because the error avoidant group completed the tasks more quickly (because each step was prescribed), they were asked to repeat the process working through the tasks.

The error training group (n = 15) was not given any detailed instruction on how to proceed. To equate the two sets of conditions, subjects in both groups were given the same tasks. However, the error training subjects received a leaflet with a short description of the commands needed to solve the tasks. Through the use of this procedure, there were ample chances for subjects to make errors. To support the possibility of learning from errors and to counter the emotional and frustrating quality of errors, a set of three heuristics was explicitly presented on a poster: "Errors are a natural part of learning. They point out what you can still learn!" "There is always a way to leave an error situation!" and "Look at the screen so that you can see what is changing there!"

After training, all subjects were tested with three identical performance tasks. The first task repeated what was learned in training; computing a correlation. The other tasks increased in difficulty, with the second requiring the calculation of a t test and the third requiring a two-factor analysis of covariance, using a regression method.

**Instruments**

During training and performance testing the computer monitor was videorecorded. The recording was used to get the following ratings:

1. **Performance rating:** Subjects' performance was rated for each of the three tasks on the following three criteria: Was the problem solved? How many deviations from a correct way were performed? How long did the problem solving take? Although there is some danger of rater bias, much of the rating consists of hard data like solving the task versus not solving it (scale of 1–5: inter-rater correlations were \( r = .63, p < .01 \) for the correlation task, \( r = .79, p < .01 \) for the t-test task, and \( r = .66, p < .01 \) for the analysis of covariance task for a subset of 15 subjects).

2. **Exploratory behavior:** The frequency and intensity of subjects' exploratory behavior during the training was rated on a scale of 1–5 (inter-rater agree-

---

**RESULTS AND DISCUSSION**

**Mean Differences Between Experimental Groups**

Table 1 presents the mean performance differences between the two training groups. Essentially, these data replicated the findings of Frese et al. (1991). The easy task produced no significant differences between the error avoidant and the error training groups. However, as predicted, the more complex tasks showed significant performance differences between the two training conditions; in each case, the error training group's competence was higher. Thus, it was again shown that error training is superior to error avoidant training and that this finding can be generalized to other areas of human-computer interactions and is not specific to word processing, as in the Frese et al. (1991) study.

**Relevance of Exploratory Behavior**

The hypothesis here was that subjects demonstrating more exploratory behavior would also perform better. Again, this hypothesis should be more pronounced for the difficult tasks than for the easy ones. Because exploratory behavior had to be operationalized differently across the two groups, the correlations between exploratory behavior and performance were calculated separately for each (see Table 2). The hypothesis was clearly supported for both groups.\(^1\) In four out of six cases, there was a significant correlation between exploratory behavior and subsequent performance.

This finding is particularly interesting in the case of the error avoidant group, because it was precisely the noncompliant subjects who resisted our error

---

\(^1\)In Table 2, the error avoidant group showed an increase in correlation with the difficulty of the tasks, the highest correlation involving the most difficult task. This increase was, however, not apparent in the error training group (there is even a nonsignificant drop in correlations from the easy to the more difficult tasks).
avoidant training procedure and explored anyway, showing the highest performance. This means that exploration is an effective mediator of good performance in computer skills training. It is interesting to see that the same relationship holds in the error training group. Thus, exploratory behavior is an effective mediator for the error training procedure as well.

The question can be asked whether the differences between the two experimental groups shown in Table 1 were actually due to exploration. Because a direct comparison of the exploration across the two groups cannot be made, an indirect way to look at this question is shown by the presentation of the results in Table 3. Here subjects were divided into those who explored a lot and those who explored little, separately for the two experimental groups. One can recognize a ranking from little to high exploration and from error avoidance training and error training. An overall analysis of variance was significant for the average ($F(3,26) = 4.76, p < .01$) and difficult tasks ($F(3,26) = 5.02, p < .01$). The first subgroup was clearly the worst. The second subgroup (more exploratory behavior and error avoidance training) showed a similar performance to the third subgroup (little exploration in error training condition). The fourth subgroup (high exploratory behavior and error training) was the best one. Thus, exploration reduces the differences between the two training groups, but there is still some superiority in the error training group that is not accounted for by the exploratory behavior. These data suggest that exploration is an effective mediator of the error training group but that there may be additional reasons for the superiority of the error training group than exploration.

### Table 1. Mean Differences in Performance Between the Two Treatment Groups

<table>
<thead>
<tr>
<th></th>
<th>Error Training ($n = 15$)</th>
<th>Error Avoidant Training ($n = 15$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Performance task: easy</td>
<td>3.32</td>
<td>1.29</td>
</tr>
<tr>
<td>Performance task: average</td>
<td>3.87</td>
<td>1.25</td>
</tr>
<tr>
<td>Performance task: difficult</td>
<td>3.93</td>
<td>1.16</td>
</tr>
</tbody>
</table>

** $p < .01$.

### Table 2. Spearman Correlation Coefficients Between Performance and Exploratory Behavior

<table>
<thead>
<tr>
<th></th>
<th>Error Training ($n = 15$)</th>
<th>Error Avoidant Training ($n = 15$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance task: easy</td>
<td>.63**</td>
<td>.37</td>
</tr>
<tr>
<td>Performance task: average</td>
<td>.54*</td>
<td>.38</td>
</tr>
<tr>
<td>Performance task: difficult</td>
<td>.44*</td>
<td>.59**</td>
</tr>
</tbody>
</table>

* $p < .05$.

** $p < .01$.

### Discussion

The data replicate and extend the finding of Frese et al. (1991) that error training leads to higher performance in human-computer interaction. Because these results have already been replicated before (Irmer et al., 1991; Thiemann, 1990), this is not surprising per se. However, the issue of the generalizability of the results is important here. Because the SPSS/PC program is an example of a poor ergonomically designed system with unhelpful error messages that do not appear immediately after the error has been made, and little help in case somebody gets lost in the system, the results show that the error training is useful not only in ergonomically sophisticated programs such as word processing but also in less well developed systems from a software ergonomic point of view.

The most important finding of this experiment is the function of exploratory behavior in training. For both groups, exploratory behavior predicted performance. Thus, this result may explain why some people learn even from "bad" training programs. Those people who did not comply with this Skinnerian type of training procedure did much better than those who complied. Possibly, many trainers misattribute the fact that something is learned to their own training skills rather than to the fact that people are resilient and resort to means that go against the training procedure used in order to learn well in spite of it. Thus, exploration is useful with two rather different training procedures.

The experiment shows that exploratory behavior is also an effective mediator in error training, although it does not explain away the differences between the error training and error avoidance groups. The results may also suggest that errors can increase learning via the mechanism of instigating exploration. An error leads people to wonder where they are in the system and what kinds of commands can be used at this point. This then increases further exploration. However, within this study, the hypothesis that errors instigate exploration cannot be proven. One would have to do microanalytic studies in which the exact exploratory behavior after each error is studied in detail.
REFERENCES


---

An Empirical Evaluation of Knowledge, Cognitive Style, and Structure upon the Performance of a Hypertext Task

Ralph D. Korthauer
United States Air Force

Richard J. Koubek
Purdue University

In an effort to reduce navigation difficulties, recent research suggests that user characteristics need to be considered when designing a hypertext document. This study evaluated the effects of an individual's cognitive characteristics in conjunction with the nature of the hypertext task. Subjects of varying levels of knowledge and cognitive style, specifically field independence/field dependence, utilized a hypertext document to answer explicitly and inherently structured questions. Among other findings, experienced field-dependent subjects' performance was less than that of experienced field-independent subjects', specifically when explicit structure was provided. The results suggest that the existing knowledge structure of experienced field-dependent subjects may conflict with explicitly structured hypertext. Implications for the appropriate design of hypertext systems are discussed.

This work was supported by the Air Force Institute of Technology and the National Science Foundation under grant number DDM 9215703. The research was performed while the first author was at Purdue University.

The authors are indebted to Dr. Stephan Konz for his input to this project and Dr. Mark Leh, who constructed the hypertext document.

Correspondence and requests for reprints should be sent to Richard J. Koubek, School Industrial Engineering, Purdue University, 1207 Grissom Hall, West Lafayette, IN 47907.
REFERENCES


An Empirical Evaluation of Knowledge, Cognitive Style, and Structure upon the Performance of a Hypertext Task

Ralph D. Korthauer
United States Air Force

Richard J. Koubek
Purdue University

In an effort to reduce navigation difficulties, recent research suggests that user characteristics need to be considered when designing a hypertext document. This study evaluated the effects of an individual's cognitive characteristics in conjunction with the nature of the hypertext task. Subjects of varying levels of knowledge and cognitive style, specifically field independence/field dependence, utilized a hypertext document to answer explicitly and inherently structured questions. Among other findings, experienced field-dependent subjects' performance was less than that of experienced field-independent subjects, specifically when explicit structure was provided. The results suggest that the existing knowledge structure of experienced field-dependent subjects may conflict with explicitly structured hypertext. Implications for the appropriate design of hypertext systems are discussed.

This work was supported by the Air Force Institute of Technology and the National Science Foundation under grant number DDM 9215703. The research was performed while the first author was at Purdue University.

The authors are indebted to Dr. Stephan Konz for his input to this project and Dr. Mark Lehto, who constructed the hypertext document.

Correspondence and requests for reprints should be sent to Richard J. Koubek, School of Industrial Engineering, Purdue University, 1287 Grissom Hall, West Lafayette, IN 47907.

REFERENCES


In an effort to reduce navigation difficulties, recent research suggests that user characteristics need to be considered when designing a hypertext document. This study evaluated the effects of an individual's cognitive characteristics in conjunction with the nature of the hypertext task. Subjects of varying levels of knowledge and cognitive style, specifically field independence/field dependence, utilized a hypertext document to answer explicitly and inherently structured questions. Among other findings, experienced field-dependent subjects' performance was less than that of experienced field-independent subjects, specifically when explicit structure was provided. The results suggest that the existing knowledge structure of experienced field-dependent subjects may conflict with explicitly structured hypertext. Implications for the appropriate design of hypertext systems are discussed.

This work was supported by the Air Force Institute of Technology and the National Science Foundation under grant number D815703. The research was performed while the first author was at Purdue University.

The authors are indebted to Dr. Stephen Konz for his input to this project and Dr. Mark Lehri who constructed the hypertext document.

Correspondence and requests for reprints should be sent to Richard J. Koubek, School of Industrial Engineering, Purdue University, 1287 Grissom Hall, West Lafayette, IN 47907.

International Journal of Human-Computer Interaction 6(4) 373-390 (1994)