

# Effectiveness of Error Management Training: A Meta-Analysis

Nina Keith and Michael Frese  
Justus-Liebig University of Giessen

Error management training (EMT) is a training method that involves active exploration as well as explicit encouragement for learners to make errors during training and to learn from them. Past evaluation studies, which compared skill-based training outcomes of EMT with those of proceduralized error-avoidant training or of exploratory training without error encouragement, have yielded considerable variation in effect sizes. The present meta-analysis compiles the results of the existing studies and seeks to explain this variation. Although the mean effect of EMT across all 24 identified studies ( $N = 2,183$ ) was positive and significant (Cohen's  $d = 0.44$ ), there were several moderators. Moderator analyses showed effect sizes to be larger (a) for posttraining transfer ( $d = 0.56$ ) than for within-training performance and (b) for performance tasks that were structurally distinct (adaptive transfer;  $d = 0.80$ ) than for tasks that were similar to training (analogical transfer). In addition, both active exploration and error encouragement were identified as effective elements in EMT. Results suggest that EMT may be better suited than error-avoidant training methods for promotion of transfer to novel tasks.

*Keywords:* errors, training, training evaluation, adaptive transfer

Errors at work are a nuisance. Errors interrupt the work flow; error correction can be time consuming and frustrating, and some workplace errors have severe consequences for individuals and for organizations. It is therefore not surprising that people usually prefer to avoid errors in the first place. Consistent with this approach, many scholars in the area of learning and training have taken a negative view of errors. A famous example is Skinner (1953), who equated errors with punishment that can inhibit behavior but that does not contribute to learning. Similarly, Bandura (1986) viewed errors as detrimental to learning and promoted a guided and error-free learning environment. In his classical monograph on social-cognitive theory, he stated that “without informative guidance, much of one’s efforts would be expended on costly errors and needless toil” (Bandura, 1986, p. 47). The present research deals with a training method that, in contrast to these approaches, takes an explicitly positive view of errors during training. This training method, which is called error management training (EMT), is based on the assumption that errors are a natural by-product of active learning: As learners actively explore the environment, errors will inevitably occur. Furthermore, errors can have an informative function for the learner, as they pinpoint where knowledge and skills need further improvement (Ivancic & Hesketh, 1995/1996). Therefore, participants in EMT are explicitly encouraged to make errors during training and to learn from them.

In early studies (e.g., Frese et al., 1991), EMT was applied to teach software skills. To determine training effectiveness, these

studies compared skill-based training outcomes of EMT with those of alternative training methods. Most of these alternative methods were proceduralized training methods, which mimic conventional tutorials that adopt a negative attitude toward errors: Detailed step-by-step instructions on correct task solutions were provided to prevent participants from making errors. Other studies compared EMT with exploratory training methods that contained no more task information than did the EMT condition and that lacked the explicit encouragement and positive framing of errors during practice or even gave instructions to avoid errors. Early studies consistently reported EMT to be effective in terms of posttraining outcomes (i.e., scores on tests given to participants after training). For example, four studies (described in Frese, 1995) reported positive and large effect sizes (Cohen’s  $d$  of about 1) in favor of EMT, which indicated that EMT participants on average outperformed those of the comparison training by one standard deviation. Later studies (e.g., Nordstrom, Wendland, & Williams, 1998; Wood, Kakebeeke, Debowski, & Frese, 2000) replicated this effect with slightly lower effect sizes. Yet, other studies found opposite results. For example, a study that used an electronic search task (Debowski, Wood, & Bandura, 2001) found that a proceduralized training group performed better than did the EMT group by three quarters of a standard deviation ( $\eta^2$  value from analysis of covariance [ANOVA] = .13, which corresponds to about  $d = -0.75$ ). Similarly, in a study that used a decision-making task (Gully, Payne, Koles, & Whiteman, 2002), a training group that received explicit instructions to avoid errors during training performed significantly better than did an EMT group, with a medium effect size ( $d = -0.50$ ).

Why would studies that, at least at first glance, draw on the same theoretical background and use similar designs come to such opposite results? Can systematic sources be identified that account for these differences, and, if so, how are they related to prevalent issues in learning and training research? The present research seeks to answer these questions. We propose that EMT can lead to

---

Nina Keith and Michael Frese, Department of Work and Organizational Psychology, Justus-Liebig University of Giessen, Giessen, Germany.

We thank Leonore Schulze for her assistance with the literature search and coding.

Correspondence concerning this article should be addressed to Nina Keith, Justus-Liebig University of Giessen, Department of Work and Organizational Psychology, Otto-Behagel Street 10 F, 35394 Giessen, Germany. E-mail: nina.keith@psychol.uni-giessen.de

better training outcomes than can exploratory or proceduralized training methods that do not utilize errors at all or that even emphasize error avoidance. Yet, depending on the particular features of how the training is conducted and evaluated, the effectiveness of EMT—as reflected in the direction and magnitude of the study’s effect size—will vary. To this end, we have applied meta-analytical techniques to existing studies that compare EMT with alternative training methods (i.e., purely exploratory or proceduralized training methods) and have systematically searched for moderators. In this article, we describe the theoretical background and typical study design of existing studies on EMT. Then, we address some issues prevalent in the learning and training literature and develop hypotheses about factors that moderate the effectiveness of EMT.

### Theory and Design of EMT

There are two characteristics of EMT that distinguish it from alternative training methods, such as proceduralized training and purely exploratory training. First, participants are given only minimal guidance and otherwise are encouraged to actively explore and experiment on their own. In addition, EMT creates a learning environment in which errors are likely to occur. For example, in software training, participants are given only some basic information on the structure and functions of the software. They are then asked to work independently on difficult training tasks without any additional information about how to solve the tasks—a procedure that almost inevitably leads to many errors. In respect to this minimal guidance, EMT is similar to exploratory training, but it differs from proceduralized training methods that seek to minimize errors by providing detailed, step-by-step instructions on correct task solutions. Second, EMT involves explicit encouragement of errors. Participants are given so-called error management instructions, which are brief instructions that tell participants to expect errors while they work on the training tasks and that emphasize the positive informational feedback of errors for learning (Frese et al., 1991). The core idea of these instructions is then summarized in a positive statement, such as “The more errors you make, the more you learn!” or “You have made an error? Great! Because now you can learn something new!” During the training session, the trainer repeats these statements and reminds participants to reflect on errors whenever they happen but provides no further assistance when an error occurs. This emphasis on and positive framing of errors during training is not present in purely exploratory training or in proceduralized training methods, which usually do not mention the issue of errors at all or may even instruct participants to avoid errors during training.

The theoretical foundation of EMT is action theory, which describes action-oriented mental models as the basis of work-related actions (Frese & Zapf, 1994; Hacker, 1998). For example, an individual’s action-oriented mental model of a photocopier machine may entail various aspects of how the machine works (e.g., how the original copy is “read” in, how the copy paper “travels” within the machine), which enables the individual to operate the machine effectively as well as to react to potential problems (e.g., to place the original copy in the right space, to fix a paper jam if it occurs). The more adequate the mental model, the more successful the actions will be, and adequate mental models are best acquired by actively dealing with the subject matter (e.g., by

actually operating the photocopier machine rather than by reading the manual). In this context, errors serve an important feedback function, because they indicate where one’s mental model is not adequately developed and thereby encourage its correction (Frese, 1995).

The view of errors as feedback is consistent with other theories that stress the importance of feedback for learning (Kluger & DeNisi, 1996; Latham & Locke, 1991), but it goes beyond regarding errors as negative feedback that indicates nonachievement of a particular goal. Rather, learners are encouraged to use errors as a basis to think ahead and to try out something new. This focus on informative aspects of errors is a distinctive feature of the error management approach (Frese, 1995; Ivancic & Hesketh, 1995/1996). The emphasis on active exploration as the primary method of learning is consistent with other active learning approaches (Bruner, 1966; Greif & Keller, 1990; Hesketh & Ivancic, 2002), but it is in contrast to approaches that stress guidance and correct behaviors during practice. In social-cognitive theory, for example, which is the theoretical basis of behavior modeling training, active exploration and errors are viewed as needless and time consuming. According to Bandura (1986), learners should be “spared the costs and pain of faulty effort” (p. 47) and should instead receive guidance that leads to flawless behavior. Feedback should focus on positive aspects of the learner’s behavior and should be given in the form of positive social reinforcement, that is, praise for correct execution of tasks (Taylor, Russ-Eft, & Chan, 2005).

### Prior Research on EMT and Aims of the Present Study

To evaluate the effectiveness of EMT, most studies used a design that compared EMT with an alternative training method, such as proceduralized or purely exploratory training, in terms of skill-based learning criteria (Kirkpatrick, 1987; Kraiger, Ford, & Salas, 1993). Examples of these criteria include number of tasks solved successfully (Chillarege, Nordstrom, & Williams, 2003; Heimbeck, Frese, Sonnentag, & Keith, 2003; Nordstrom et al., 1998; Wood et al., 2000); correctness, efficiency, and speed of solutions in difficult tasks (Dormann & Frese, 1994; Frese et al., 1991); and number of errors in transfer tasks (Ivancic & Hesketh, 2000). The majority of studies, particularly the earlier ones, were conducted in the area of software skills (Frese, 1995); other studies investigated decision-making tasks (e.g., Gully et al., 2002) or used EMT in driver training (Ivancic & Hesketh, 2000). The major aim in the earlier studies was to assess the overall effectiveness of EMT compared with proceduralized error-avoidant training methods. More recent studies have focused on interactions of personal characteristics and training method (e.g., Gully et al., 2002; Heimbeck et al., 2003) or on emotional or cognitive processes that potentially explain the effectiveness of EMT (e.g., Chillarege et al., 2003; Nordstrom et al., 1998; Wood et al., 2000). In a recent study, Keith and Frese (2005) found empirical support for the notion that EMT increases participants’ tendency to use two self-regulatory skills: Participants learn to exert emotion control aimed at reducing negative emotional reactions to errors and setbacks (Kanfer, Ackerman, & Heggstad, 1996), and they engage in metacognitive activities that involve planning, monitoring, and evaluating one’s progress during task completion and revision of strategies (Brown, Bransford, Ferrara, & Campione, 1983). Such metacognitive activities are instigated because “errors prompt

learners to stop and think about the causes of the error” (Ivancic & Hesketh, 2000, p. 1968) and to experiment with different solutions.

Given the existing empirical evidence and the theoretical propositions that describe how making errors can potentially be fruitful for learning and performance, EMT may be expected to be generally effective compared with alternative training methods that do not encourage errors during training. As we have pointed out, however, considerable variation exists among effect sizes from studies that evaluate the effectiveness of EMT. The major aim in the present research is to identify variables that account for this variation—variables that, technically speaking, moderate the effectiveness of EMT. In the following pages, we derive several moderator hypotheses, based on prevalent issues in the learning and training literature, that we suggest may be particularly relevant for EMT (see Figure 1).

### Within-Training Performance Versus Posttraining Transfer Performance

In classical comparative-training studies, two (or more) groups of participants practice the same tasks under different training conditions (e.g., different instructions, different frequencies of feedback; Hesketh, 1997). Training effectiveness is then ascertained on the basis of participants’ performance on a separate task, which is given to them after the actual training and which is the same across groups. In other words, the independent variable is manipulated in the training phase, during which the skills are acquired; the dependent variable is collected in a posttraining transfer phase, during which the learned skills are applied to separate tasks. This conceptual and operational distinction is essential; it ensures that potential performance differences between groups do not vanish, once the particular manipulation is removed, and that they can really be attributed to relatively permanent changes, which “qualify for the label *learning effects*” (R. A. Schmidt & Bjork, 1992, p. 208). A second reason why this distinction is important is the phenomenon that manipulations which appear to boost immediate performance during training may be relatively useless for posttraining transfer performance and vice versa. A prominent example is a study by Shea and Morgan (1979) in which participants learned three movement tasks, either in a blocked practice or in a randomized order. During training, the blocked-practice group per-

formed better than did the random-order group. On a posttraining transfer task, however, the pattern was reversed; participants who had learned under the random-order condition outperformed those who had learned under the blocked-practice condition, particularly if the transfer task was given in random order.

From a practical perspective, these results (and similar results found in domains other than simple movements; cf. R. A. Schmidt & Bjork, 1992) imply (a) that conclusions about training effectiveness can be misleading if they are based on skill-based measures that are assessed within the training phase and (b) that training designs should generally include an additional posttraining transfer phase for evaluation. Conceptually, the results imply that introducing difficulties during training may enhance transfer, at least if these difficulties elicit psychological processes that are also useful during transfer (Hesketh, 1997; Salas & Cannon-Bowers, 2001). EMT, which gives learners ample opportunities to make errors during training, is an example of such a training method: It introduces difficulties during training that can be beneficial for transfer, because the processes elicited during training, such as emotion control and metacognition, promote learning (Ivancic & Hesketh, 1995/1996; Keith & Frese, 2005). Accordingly, in EMT, compared with training methods that do not encourage errors during training, immediate performance within the training phase may not differ or may even be depressed, as participants make errors, explore, and sometimes arrive at suboptimal solutions. Positive effects of EMT can be expected for performance in a posttraining transfer phase, in which participants are aware that their performance is being evaluated and that errors are no longer encouraged (cf. Wood et al., 2000).

*Hypothesis 1:* EMT leads to better posttraining transfer performance but not to better within-training performance than do proceduralized or exploratory training methods that do not encourage errors during training.

In operational terms, Hypothesis 1 implies that the type of evaluation outcome moderates the effectiveness of EMT compared with that of alternative training methods. Studies that use posttraining transfer performance as the criterion will yield larger effect sizes than will studies that use within-training performance; for the former studies only, the mean effect size will be significant (see Figure 1).

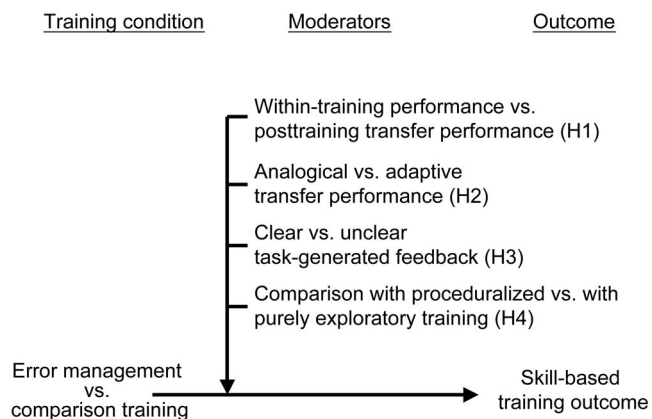


Figure 1. Meta-analytical hypotheses about factors moderating the effectiveness of error management training. H = hypothesis.

### Analogical Versus Adaptive Transfer Tasks

The previous section deals with the issue of posttraining transfer performance (as opposed to within-training performance) in general. Yet, it can be useful to distinguish between different types of transfer when one is evaluating training effectiveness. One such distinction, which is based on the similarity of training and transfer tasks, is between analogical and adaptive transfer (Ivancic & Hesketh, 2000; see Barnett & Ceci, 2002, for a similar distinction between near and far transfer). Transfer implies that knowledge and skills are “transferred from one task or job to another” (Hesketh, 1997, p. 318). Analogical transfer refers to situations where problem solutions of the transfer tasks are familiar or analogous to those of the training tasks. Adaptive transfer comprises “using one’s existing knowledge base to change a learned procedure, or to generate a solution to a completely new problem” (Ivancic &

Hesketh, 2000, p. 1968). Adaptive transfer implies that rote application of a procedure learned in training is not sufficient and that the problem at hand is structurally different from those encountered during training and requires the learner to modify the learned procedures (Ivancic & Hesketh, 1995/1996).

The goal of a particular training program can be to promote analogical transfer, for example, when a limited and clear-cut behavioral repertoire is to be performed on a job that can, in principle, be taught within the allotted training time. In cases in which not all potential work-related problems and their solutions can be taught, however, the training goal may be to promote adaptive transfer, that is, to enable participants to develop new solutions to structurally novel problems by using and adapting the skills they acquired during training. Training researchers have suggested that explicitly allowing and encouraging errors to occur during training, as is done in EMT, may be one means for promotion of adaptive transfer (e.g., Ivancic & Hesketh, 1995/1996; Smith, Ford, & Kozlowski, 1997). When confronted with errors during training, participants may engage in mindful processing (Salomon & Perkins, 1989), such as metacognition (Keith & Frese, 2005), and thereby gain knowledge and acquire skills that are particularly useful in solving structurally distinct adaptive transfer tasks (Ivancic & Hesketh, 2000). EMT may promote analogical transfer as well, because errors during training instigate attention, which in turn facilitates later retrieval of similar problems and their solutions (Ivancic & Hesketh, 1995/1996, 2000). Yet, to promote analogical transfer, other training methods that emphasize error-free learning and correct procedures for a particular task may be equally effective, as they directly teach the required procedures that are then merely applied to the similar transfer task. As a consequence, the advantage of EMT in comparison with such training methods can be expected to be smaller for analogical than for adaptive transfer tasks.

*Hypothesis 2:* The effectiveness of EMT in comparison with that of proceduralized or exploratory training methods that do not encourage errors during training will be more pronounced for adaptive than for analogical transfer tasks.

In operational terms, Hypothesis 2 implies that the type of transfer tasks used in studies moderates the effectiveness of EMT compared with that of alternative training methods. Studies that use adaptive transfer tasks will yield larger effect sizes than will studies that use analogical transfer tasks; for both groups of studies, the mean effect size will be significant (see Figure 1).

### Task-Generated Feedback

As outlined above, action theory views errors during training as valuable pieces of information, because they serve as feedback for one's actions and can point out what aspects of one's knowledge need further correction and refinement (Frese & Zapf, 1994). In general, feedback permits an individual to judge the extent to which he or she has achieved the goal or standard (Carver & Scheier, 1998; Frese & Zapf, 1994; Hacker, 1998; Ilgen, Fisher, & Taylor, 1979; Latham & Locke, 1991; Sonnentag, 1998). Errors are one form of negative feedback that indicates a deviation between the goal or standard and the current state (Frese & Zapf, 1994). In addition to the judgment about the current state, one can

use errors and feedback retrospectively to evaluate the effectiveness of one's previous strategies (Neubert, 1998) and, based on this evaluation, to adjust one's strategies accordingly. For this positive function of errors to take effect, however, it is a precondition that errors can be detected at all. Given that EMT provides little external guidance—participants work independently during training and without constant monitoring by a trainer who informs about the correctness of their actions—the task itself needs to provide clearly interpretable and informative feedback if EMT is to be effective (e.g., noticeable visual changes on the display that appear in response to the user's actions are one type of informative feedback). Stated differently, on tasks that provide only feedback that cannot be readily interpreted, EMT may be less suitable compared with alternative training methods that give external informative guidance (Debowski et al., 2001).

*Hypothesis 3:* The effectiveness of EMT in comparison with that of proceduralized or exploratory training methods that do not encourage errors during training is limited to tasks that provide clear feedback.

In operational terms, Hypothesis 3 implies that clarity of the task-generated feedback moderates the effectiveness of EMT compared with that of alternative training methods. Studies that use clear-feedback tasks will yield larger effect sizes than will studies that use tasks with unclear feedback; for those studies that use clear-feedback tasks, the mean effect size will be significant (see Figure 1).

### Effective Elements of EMT: Active Exploration and Error Management Instructions

There are two elements in EMT that distinguish it from alternative training methods: the element of active exploration (i.e., participants receive little external guidance and explore the tasks independently) and the element of error encouragement (i.e., participants receive error management instructions that frame errors positively and encourage errors during training). As stated above, error management theory assumes both elements to be effective in EMT: Active exploration is important, because adequate mental models are best acquired by direct action. A positive view of errors, as is conveyed in error management instructions, is essential if learning is to occur, because "developing an error tolerant attitude . . . maximize[s] the informational value of errors" (Ivancic & Hesketh, 1995/1996, p. 115). If errors are not tolerated but are viewed negatively, participants will likely be frustrated by errors and will refrain from further exploration, with a resulting decrease in learning opportunities (cf. van Dyck, Frese, Baer, & Sonnentag, 2005).

Recent studies have—and justifiably so—criticized that studies comparing EMT with proceduralized error-avoidant training, which differs from EMT in that it does not include either exploration or encouragement of errors, confound these two elements. As a result, observed differences in training outcomes cannot be unequivocally attributed to either element (e.g., Bell & Kozlowski, 2007; Gully et al., 2002). There are, however, a few attempts in the literature on EMT to dissociate the two elements by varying them both. For example, Heimbeck et al. (2003) found that EMT (i.e., active exploration with error management instructions) led to

better training outcomes than did either proceduralized training (i.e., no exploration, no error management instructions) or purely exploratory training (i.e., active exploration only, no error management instructions). In addition, there are some EMT studies that do not vary the exploration element but only the element of error encouragement (i.e., error management instructions vs. no such instructions). If we assume both elements of EMT, active exploration and error encouragement, to be additively effective elements, the following meta-analytical pattern should emerge: The comparisons of EMT with proceduralized training and with exploratory training should yield significant differences. Moreover, the comparison with proceduralized training should yield larger differences than should the comparison with exploratory training (because the combination of two elements, exploration and error encouragement, should yield larger differences than the difference yielded by one element of exploration).

*Hypothesis 4:* Both active exploration and error management instructions are effective elements in EMT.

In operational terms, Hypothesis 4 implies that the type of comparison training method moderates the effectiveness of EMT. Studies that compare EMT with proceduralized error-avoidant training will yield larger effect sizes than will studies that compare EMT with exploratory training (see Figure 1). At the same time, the comparison of EMT with exploratory training will still yield significant mean effect sizes.

## Method

### *Pool of Primary Studies*

To identify empirical studies that investigated the effectiveness of EMT, we conducted an electronic search in relevant databases (PsycINFO, Social Sciences Citation Index, and the German database Psycdex), supplemented by manual searches of conference programs and reference lists of identified studies. We also contacted authors of published papers and other researchers in the field. (The initial search was conducted in spring 2004; the database search was updated in fall 2006.) In line with the theory of EMT outlined above, training methods were considered to represent EMT if they met the following two criteria: First, the training involved active exploration, in that participants were not guided to correct task solutions but worked independently to find solutions on their own. Second, the training explicitly encouraged making errors by providing participants with error management instructions that stressed the positive function of errors for learning. The search yielded 24 studies (overall  $N = 2,183$ ) that met these criteria. Another 3 studies were identified that explicitly referred to the theory of EMT but that did not operationalize EMT according to the aforementioned criteria (for example, participants did not actively explore and make errors themselves but were presented with potential errors selected by the researchers, or participants did not receive error management instructions; Ivancic & Hesketh, 2000; Joung, Hesketh, & Neal, 2006; Lorenzet, Salas, & Tannenbaum, 2005). To retain a study pool that was sufficiently homogeneous with regard to the training method examined, we omitted these 3 studies from the present meta-analysis (cf. Oswald & McCloy, 2003).

All 24 studies evaluated training effectiveness by comparing EMT with an alternative training (i.e., we did not identify any studies that used a no-training control group as comparison). Thus, all effect sizes compiled in the present meta-analysis refer to relative effectiveness (i.e., compared with an alternative training) rather than to absolute effectiveness of EMT (i.e., compared with no training at all). Of the 24 studies, 11 were unpublished (e.g., master's theses/dissertations, technical reports, or manuscripts in preparation). All studies but 1 were experimental laboratory studies, with participants or small groups of participants randomly assigned to training conditions. The majority of the trainings taught a new software ( $k = 18$ ) or electronic search of databases ( $k = 3$ ). The remaining 3 trainings used a computerized decision-making task. Participant samples of 6 studies consisted of employees recruited in the community; the remaining studies used samples of university students. In 2 studies, participant dyads (rather than individuals) worked together and were the unit of analysis.

If studies compared one or more EMT conditions with more than one other training condition, we calculated the mean effect size for use in further analyses. Similarly, if studies assessed multiple training outcomes, we included the mean effect size. If the multiple outcomes assessed were not similar but referred to tasks of different levels of difficulty, we included only the most difficult task, because this is the most relevant variable from a theoretical point of view. It is recommended that researchers use only one effect size per study to avoid statistical dependencies (Hedges & Olkin, 1985; Hunter & Schmidt, 1990; Lipsey & Wilson, 2001; Rosenthal, 1991). Brief descriptions of the 24 studies as well as their effect sizes are listed in Table 1.

### *Coding of Study Characteristics and Interrater Agreement*

Corresponding to our research hypotheses, we coded four study characteristics as potential moderators. All but one of the 24 studies were coded independently by a second rater. Interrater agreement (Cohen's kappa) was good to excellent according to Fleiss (1981), with coefficients ranging from .65, for clarity of task feedback, to 1, for type of comparison training (.83 for type of evaluation outcome; .91 for type of transfer tasks). Cases in which initial codings of the two raters differed were resolved by discussion.

The dichotomous variable type of evaluation outcome was designed to describe whether the study design included a separate posttraining transfer phase for performance evaluation, in which participants were aware that their performance was being evaluated (posttraining transfer performance), or whether no such separate phase existed. For example, if the performance score in one of several training trials (during which, in the case of EMT, errors were still encouraged) served as the criterion performance in the study, this was coded as "within-training performance."

The dichotomous variable type of transfer task was designed to capture whether the criterion tasks were structurally similar to the tasks that participants had worked on during training (analogical transfer) or whether the tasks required the development of new solutions (adaptive transfer; Ivancic & Hesketh, 2000). The critical dimension was task distinctiveness rather than task difficulty (Frese & Zapf, 1994; Keith & Frese, 2005). For example, if participants were tested on the same tasks as in training but under greater time pressure, these tasks might be more difficult but not be structurally distinct, which would indicate analogical transfer.

Table 1  
*Brief Descriptions and Statistics of Included Studies*

Study	Training content	Alternative training condition <sup>a</sup>	<i>d</i> (SE)
Bell & Kozlowski (2007) <sup>b</sup>	PC-based decision-making simulation	Proceduralized training with error encouragement or avoidance instructions (4). Exploratory training with error avoidance instructions (2).	0.33 (0.12)
Chillarege et al. (2003) <sup>b</sup>	Software: Word processor	Proceduralized training with error avoidance instructions (1).	1.35 (0.27)
Debowski et al. (2001)	Electronic database search	Proceduralized (guided) training without error-related instructions; immediate error correction by trainer (1).	-0.73 (0.30)
Dormann & Frese (1994)	Software: Statistical package	Proceduralized training without error-related instructions; immediate error correction by trainer (1).	1.08 (0.39)
Frese et al. (1991)	Software: Word processor	Proceduralized training without error-related instructions; immediate error-correction by trainer (1).	0.99 (0.45)
Granados (2000)	Software: Presentation	Exploratory training without error-related instructions (1).	0.30 (0.32)
Greif & Janikowski (1987)	Software: Word processor	Proceduralized training/tutorial (1).	1.16 (0.62)
Gully et al. (2002)	PC-based decision-making simulation	Exploratory training with error avoidance instructions (1). Exploratory training without error-related instructions (1).	-0.44 (0.16)
Heimbeck et al. (2003)	Software: Spreadsheet	Proceduralized training with error avoidance instructions; immediate error correction by trainer (1). Exploratory training without error-related instructions (1).	0.72 (0.23)
Heinbokel (1990)	Software: Word processor	Proceduralized training without error-related instructions (1).	1.06 (0.53)
Irmer et al. (1991)	Software: Word processor	Proceduralized training without error-related instructions (standard training of software training school); immediate error correction by trainer (1).	0.95 (0.34)
Ivancic (1998), Study 1 <sup>b</sup>	Software: E-mail	Proceduralized training with error encouragement or error avoidance instructions (4). Exploratory training with error avoidance instructions (2).	0.22 (0.37)
Ivancic (1998) <sup>1</sup> , Study 3 <sup>b</sup>	Software: E-mail	Exploratory training without error-related instructions (2).	-0.24 (0.30)
Keith & Frese (2005)	Software: Presentation	Proceduralized training without error-related instructions (1).	0.74 (0.30)
Keith & Müller (2004)	Software: Presentation	Proceduralized training with or without error encouragement instructions (2). Exploratory training without error-related instructions (1).	0.21 (0.22)
Lazar & Norcio (2003) <sup>b</sup>	Software: Web browser	Proceduralized (traditional) training with error encouragement or avoidance instructions (4). Exploratory training with error avoidance instructions (2).	0.09 (0.14)
Nordstrom et al. (1998) <sup>b</sup>	Software: Word processor	Proceduralized training with error avoidance instructions (1).	0.53 (0.21)
Stiso & Payne (2007) <sup>b</sup>	PC-based decision-making simulation	Exploratory training without error-related instructions (1).	0.11 (0.14)
Thiemann (1990)	Software: Word processor	Proceduralized training without error-related instructions; immediate error correction by trainer (1).	1.44 (0.62)
van Dyck (2007a)	Software: Statistical package	Exploratory training with error avoidance instructions (1). Exploratory training without error-related instructions (1).	0.00 (0.36)
van Dyck (2007b), Study 1 <sup>b</sup>	Programming language	Exploratory training with error avoidance instructions (2).	0.77 (0.32)
van Dyck (2007b), Study 2 <sup>b</sup>	Programming language	Exploratory training without error-related instructions (1). Exploratory training with error avoidance instructions (2).	0.76 (0.23)
Wood et al. (2000)	Electronic database search	Exploratory training without error-related instructions (1).	0.69 (0.35)
Yorke (2006)	Software: Spreadsheet	Proceduralized training with error avoidance instructions; immediate error correction by trainer (1).	0.50 (0.14)

*Note.* Reported are Cohen's *d* effect sizes with small-sample correction for unbiased effect sizes (Hedges, 1981). The values may therefore differ slightly from those reported in the original studies. Positive effect sizes denote performance differences in favor of the error management training condition(s).  
<sup>a</sup> Proceduralized training: Participants followed detailed, step-by-step instructions. Exploratory training: Participants received basic task information and worked independently, without detailed, step-by-step instructions. Number of training conditions is shown in brackets. <sup>b</sup> Study design included additional manipulation(s) that were ineffective, irrelevant for present research questions, or both, and that were disregarded in the present meta-analysis and therefore are not described in this table (e.g., different task orders, additional manipulations unrelated to errors).

The dichotomous variable clarity of task feedback was designed to describe the feedback provided by the task. It was coded on the basis of information provided in the Method sections and, if available, appendices of the studies, such as general task descriptions (including screen shots) and descriptions of system responses to users' actions. If the feedback enabled participants to track the consequences of their actions and to detect errors, this was coded as "clear task feedback." For example, if a user inserts a table in a document and the software responds by visually displaying the table in the document, this is interpretable feedback. If there was ambiguous or no feedback or if further information (e.g., background knowledge) was required for one to understand whether an action was correct, this was coded as "unclear task feedback." For example, some statistics programs (as used in the study by Dormann & Frese, 1994) generate rather complex and comprehensive outputs that may not be readily interpretable for novice users. As another example, electronic databases provide ambiguous feedback in the sense that although the system displays several records in response to the search statements, it might be difficult for the searcher to judge the relevance of these records and, in turn, the appropriateness of his or her search strategy (Debowksi et al., 2001).

The categorical variable comparison condition described the training method that EMT was compared with in a given study. None of the comparison conditions involved explicit error encouragement (consequently, this training characteristic was not coded), but they differed in the amount of guidance. A comparison condition was coded as "proceduralized" if participants received step-by-step instructions or close personal guidance to correct task solutions during practice. A comparison condition was coded as "exploratory" if participants practiced the training tasks without

guidance, irrespective of how much guiding information they received otherwise. For example, if participants first received some information on the task but then practiced the training tasks independently, this was coded as exploratory. Some studies used more than one comparison condition, out of which some were proceduralized and others were exploratory. To account for these studies, we coded them as "both."

### Data Analytic Strategies

Meta-analytic techniques as described by Hedges and Olkin (1985) were used. Mean effect sizes were calculated with the small-sample correction formulas for unbiased effect sizes (Hedges, 1981). Moderating effects of single dichotomous variables were tested with the procedural analogue to ANOVA by Hedges (1982). All analyses were based on random or mixed effects models, which take both subject level and study-level sampling error into account. The ANOVA analogue partitions the overall variance into a portion that is explained by the independent variable (i.e., the moderator variable) and an unexplained residual portion. These two portions are represented in a  $Q$  statistic (i.e.,  $Q_{\text{between}}$  and  $Q_{\text{within}}$ ). The  $Q$  test is analogous to the  $F$  test in ANOVA or in regression analysis and can be interpreted accordingly. Thus, a significant  $Q_{\text{between}}$  statistic indicates that the moderator variable significantly explains variability of effect sizes (Lipsey & Wilson, 2001).

### Results

The overall mean effect size across all studies included in the meta-analysis was significant (see Table 2), which suggests that,

Table 2  
Overall Meta-Analytical Effect of Error Management Training, Moderator Effects, and Statistics in Subsamples

	Moderator analysis (ANOVA analogue)		Statistics in subsample							Homogeneity $Q$ ( $df$ )
			$Q_{\text{between}}$ ( $df$ )	$Q_{\text{within}}$ ( $df$ )	$k$	$n$	$d$	$SE$	$z$	
	Left	Right								
Overall effect			24	2,183	0.44	0.10	4.36**	0.27	0.61	26.46 (23)
Type of evaluation outcome	10.61** (1)	23.99 (22)								
Within-training performance			4	505	-0.15	0.20	-0.76	-0.47	0.17	1.39 (3)
Posttraining transfer performance			20	1,678	0.56	0.10	5.79**	0.40	0.73	22.60 (19)
Type of transfer	10.40** (1)	22.38 (22)								
Analogical			13	1,647	0.20	0.11	1.79*	0.02	0.39	18.31 (12)
Adaptive			11	536	0.80	0.15	5.46**	0.56	1.05	4.07 (10)
Clarity of task feedback	2.92† (1)	23.64 (22)								
Low			7	1,005	0.19	0.18	1.07	-0.10	0.49	7.15 (6)
High			17	1,178	0.56	0.13	4.60**	0.36	0.76	16.48 (16)
Comparison condition <sup>a</sup>	4.28* (1)	17.14 (17)								
Proceduralized training (no exploration, no error encouragement)			16	1,257	0.58	0.11	5.03**	0.39	0.77	18.79 (15)
Exploratory training (exploration without error encouragement)			13	1,158	0.19	0.11	1.69*	0.01	0.37	11.59 (12)

Note. Cohen's  $d$  measures effect sizes. One-tailed  $z$  tests for directional hypotheses. ANOVA = analysis of variance.

<sup>a</sup> To avoid statistical dependencies, the overall effect of comparison condition was tested with a dummy variable representing whether the comparison condition was a proceduralized training method ( $k = 11$ ) or an exploratory training method ( $k = 8$ ). To calculate statistics in subgroups, we drew two effect sizes from the studies that included both comparison conditions ( $k = 5$ ) and added them to the other studies using proceduralized or exploratory training methods for comparison, respectively. The total number of studies therefore does not add up to 24, and the total number of participants does not add up to 2,183 (as 5 studies are represented twice).

†  $p < .10$ . \*  $p < .05$ . \*\*  $p < .01$ .

overall, EMT leads to better training outcomes compared with training methods that do not encourage errors during training. The test of homogeneity was not significant, which indicates that, from a purely statistical perspective, effect size variability was not any greater than would be expected from sampling error. Yet, the homogeneity test suffers from low power for detection of true variance across studies and can often lead to the false conclusion that no moderators exist, particularly when primary studies are relatively few and are based on small samples. Therefore, we still tested our a priori theoretical moderator hypotheses (this procedure is in line with recommendations in the literature; e.g., Lipsey & Wilson, 2001; Oswald & McCloy, 2003; Rosenthal & DiMatteo, 2001; F. L. Schmidt & Hunter, 2002).

Hypothesis 1 predicted that EMT would lead to better posttraining transfer but not to better within-training performance than would proceduralized or exploratory training methods that do not encourage errors during training. This hypothesis was supported, as evaluation phase significantly affected the magnitude of the effect sizes ( $p < .01$ ; see Table 2). Studies that used a training phase for evaluation yielded a nonsignificant mean effect size ( $p = .45$ ). Studies that used a posttraining transfer phase for evaluation yielded a significant medium mean effect size ( $d = 0.56, p < .01$ ). Hypothesis 2 predicted that the effectiveness of EMT would be larger for adaptive than for analogical transfer performance. This hypothesis was supported, as type of transfer task significantly affected the magnitude of the effect sizes ( $p < .01$ ; see Table 2). Studies that used an analogical transfer task as the criterion yielded a small but significant mean effect size ( $d = 0.20, p < .05$ ), and studies that used an adaptive transfer task as the criterion yielded a significant and large mean effect size ( $d = 0.80, p < .01$ ). Hypothesis 3 predicted that the effectiveness of EMT would be limited to tasks with clear feedback. This hypothesis received only limited support: The effect of clarity of feedback on the magnitude of the effect sizes did not reach significance at the 5% level ( $p = .08$ ; see Table 2). The statistics in subsamples, however, indicated that only studies that used tasks with clear feedback yielded a significant positive effect ( $d = 0.56, p < .01$ ), whereas studies that used tasks with unclear feedback yielded a nonsignificant mean effect size ( $p = .28$ ). Hypothesis 4 predicted that both active exploration and error encouragement would be effective elements of EMT. This hypothesis was supported. First, the comparison condition significantly affected the magnitude of the effect sizes ( $p < .05$ , see Table 2). Studies that compared EMT with proceduralized training, which involves no exploration and no error encouragement, yielded higher effect sizes than did studies that compared EMT with exploratory training, which involves exploration but no error encouragement. Second, at the same time, the comparison of EMT with exploratory training yielded a small but significant mean effect size ( $d = 0.19, p < .05$ ).

## Discussion

### *Summary of Results and Implications for Theory and Practice*

The present meta-analysis compiled results from 24 studies that investigated the effectiveness of EMT. These studies compared training outcomes of EMT with those of proceduralized or exploratory training methods that did not involve explicit encouragement

of errors (i.e., relative effectiveness of EMT). The average effect across all studies was positive (Cohen's  $d = 0.44$ ), indicating that EMT leads to, on average, better training outcomes than do these alternative training methods. This result demonstrates that deliberately incorporating errors into training can be an effective means for promotion of learning—a result that is in contrast to many traditional training approaches that focus exclusively on correct behaviors and that deny any positive functions of errors during training (e.g., Bandura, 1986; Skinner, 1953).

This meta-analysis further identified several moderator variables that affected the magnitude of the effect size. First, EMT appeared to be effective only when posttraining performance and not within-training performance was considered. This result is in line with training theory and research that emphasizes the distinction between within-training and posttraining transfer performance (Goodman & Wood, 2004; Hesketh, 1997; R. A. Schmidt & Bjork, 1992). From a practical perspective, this result implies that trainers should not focus on optimizing within-training performance, which may be slowed down in EMT as participants make errors, but should keep in mind that a training method can be effective despite apparently impaired initial performance, as may be the case with EMT. Also, this result underscores empirically the call for evaluation of training effectiveness, be it of EMT or of any other training method, on the basis of posttraining outcome measures rather than of performance during training itself (Hesketh, 1997; R. A. Schmidt & Bjork, 1992).

Second, the present results showed EMT to be particularly effective when adaptive transfer rather than analogical transfer is involved. Thus, employing EMT to deliver training seems most useful when the major training goal is to transfer learned skills to novel problems that require the development of new solutions (i.e., adaptive transfer), for example, in situations in which the skills required on the job are too diverse to be covered completely during the allotted training time. When the training goal is to learn and to apply just one particular procedure, however, other training methods that involve direct instruction of this procedure may also be effective while being less time consuming and less effortful than EMT (see Ivancic & Hesketh, 1995/1996).

Third, the present meta-analysis found significant mean effect sizes not only in studies that compared EMT with proceduralized error-avoidant training (without active exploration and without error management instructions that encourage errors) but in studies that compared EMT with exploratory training (with active exploration but without error management instructions). This finding can be interpreted to the effect that both elements of EMT—namely, active exploration and explicit encouragement of errors—are effective in EMT and that any exploratory practice should be supplemented with error management instructions, given that these simple and easy-to-administer instructions can produce significant incremental effects. It would be desirable to conduct more studies that included both proceduralized and purely exploratory training in one experimental design to further examine the feasibility of this interpretation.

Finally, for one moderator, clarity of task-generated feedback, results were mixed, and it cannot be concluded that EMT was effective only when task-generated feedback was clear. This result may be due to the general usefulness of feedback for learning and performance: Although clarity of feedback may be important if EMT is to be effective, it may be just as important for the other



training methods that served as comparison training conditions for EMT. In addition, the relatively low interrater reliability for the feedback variable (Cohen's  $\kappa = .65$ ) may have contributed to the nonconclusive findings regarding this moderator.

### *Strengths, Limitations, and Directions for Future Research*

This research compiled all currently available studies that evaluated the effectiveness of EMT, defined as active exploratory training with explicit instructions that encourage errors during training. Overall, the meta-analytical results suggest that EMT is an effective training method compared with methods that do not encourage errors during training, such as purely exploratory and proceduralized training. In addition, this research illuminates the conditions under which EMT seems most promising. By the same token, it helps to resolve conflicting findings in the literature: EMT is most likely to be effective for adaptive transfer tasks that require the application of learned skills to a structurally new problem; conversely, it is likely to fare worse than other training methods if within-training performance rather than posttraining transfer is considered for evaluation.

Some cautionary remarks should be made concerning the generalizability of the present results beyond the 24 primary studies included in this meta-analysis. Almost all (21 out of 24) studies applied EMT in teaching software skills (the remaining 3 studies used decision-making tasks delivered on the computer). Clearly, more research is needed to test to what extent the present results generalize to other technical skills (i.e., operating machines other than the computer) as well as to completely different types of skills (e.g., social skills, managerial skills). A recent study by Joung et al. (2006) provides an example of how the theoretical ideas underlying EMT can be applied in a noncomputer setting. Joung et al. described a variant of EMT, in which participants do not explore and make errors themselves but in which researchers or trainers select other people's actual or potential errors, which are to be presented to and discussed with participants (i.e., vicarious EMT; due to this different operationalization of EMT, this study was not included in the present meta-analysis). Joung et al. applied this vicarious variant of EMT in a training for firefighters, who were presented with firefighting scenarios in training (i.e., firefighting stories based on real incidents). One group of trainees received scenarios that contained errors made by the incident controller (e.g., underestimation of resource requirements), whereas a second group received scenarios without errors (i.e., success stories). When presented with new scenarios, participants in the first training group outperformed members of the second group in terms of problems identified in the firefighting practices. If future research could substantiate that this kind of vicarious EMT can be similarly effective to the active form of EMT investigated in the present meta-analysis, this would open interesting opportunities for the practice of training. Vicarious EMT could be employed to incorporate errors in training if active exploration is not a viable option, for example, because of lack of adequate equipment (e.g., simulators), or, as may be the case with firefighting training, if the prospect of making errors oneself is too threatening and stressful for participants.

None of the currently available studies on EMT have actually measured on-the-job performance as a result of training. All have

employed posttraining performance tasks to evaluate training effectiveness—a research gap that certainly should be filled. For two reasons, however, it could be speculated that EMT benefits on-the-job performance. First, errors occur not only during training but on the job as well, a situation in which there is usually no trainer available to assist with error handling. EMT may be well suited to prepare for this situation, because participants learn to deal with errors independently from the very beginning (Frese, 1995; Ivancic & Hesketh, 1995/1996). Second, the present results show EMT to be particularly effective in the promotion of adaptive transfer, that is, performance on tasks that are structurally distinct from training tasks and that require the modification of a learned procedure. Both theoretical considerations (e.g., Hesketh, 1997) and empirical evidence (Keith & Frese, 2005) suggest that this is because EMT participants learn to apply metacognitive skills, which in turn prove useful when they are confronted with novel tasks not practiced during training. We suppose that these metacognitive skills, which enable participants to select and to adjust their strategies according to the task at hand, serve as generalizable skills that are beneficial for on-the-job performance (see Ford, Smith, Weissbein, Gully, & Salas, 1998; Smith et al., 1997).

Another potential limitation of the present meta-analysis is the relatively small number of studies currently available in the area of EMT training, at least compared with the number for meta-analyses that have a broader research focus (e.g., the IQ/performance relationship) or that deal with training methods that are older and more established than is EMT (e.g., behavior modeling training). The problem with smaller study pools is that they are generally associated with lower statistical power. It should be noted, however, that we found several significant moderator effects despite this low power—a result that underscores the relevance of the identified moderators for the effectiveness of EMT. In addition, although the number of available studies as well as the type of outcome measures used in the primary studies may be criticized, the design of the included studies contributes to a strength of the present meta-analysis: All but one of the studies employed a controlled experimental design with randomized assignment to training conditions. This design renders justification for causal inferences, namely, that differences in training outcomes between EMT and alternative training methods are in fact due to the training method employed and are not the result of pretraining differences or of some other uncontrolled third variable.

### References

- References marked with an asterisk indicate studies included in the meta-analysis.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, *128*, 612–637.
- \*Bell, B. S., & Kozlowski, S. W. J. (2007). *Core elements of active learning: Cognition, motivation, and emotion-control effects on self-regulated learning and adaptability*. Manuscript submitted for publication.
- Brown, A. L., Bransford, J. D., Ferrara, R. A., & Campione, J. C. (1983). Learning, remembering, and understanding. In J. H. Flavell & E. M. Markman (Eds.), *Handbook of child psychology* (Vol. 3, pp. 77–166). New York: Wiley.

- Bruner, J. S. (1966). *Toward a theory of instruction*. Cambridge, MA: Harvard University Press.
- Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior*. New York: Cambridge University Press.
- \*Chillarege, K. A., Nordstrom, C. R., & Williams, K. B. (2003). Learning from our mistakes: Error management training for mature learners. *Journal of Business and Psychology, 17*, 369–385.
- \*Debowski, S., Wood, R. E., & Bandura, A. (2001). Impact of guided exploration on self-regulatory mechanisms and information acquisition through electronic search. *Journal of Applied Psychology, 86*, 1129–1141.
- \*Dormann, T., & Frese, M. (1994). Error management training: Replication and the function of exploratory behavior. *International Journal of Human-Computer Interaction, 6*, 365–372.
- Fleiss, J. L. (1981). *Statistical methods for rates and proportions*. New York: Wiley.
- Ford, J. K., Smith, E. M., Weissbein, D. A., Gully, S. M., & Salas, E. (1998). Relationships of goal orientation, metacognitive activity, and practice strategies with learning outcomes and transfer. *Journal of Applied Psychology, 83*, 218–233.
- Frese, M. (1995). Error management in training: Conceptual and empirical results. In C. Zuccheromaglio, S. Bagnara, & S. U. Stucky (Eds.), *Organizational learning and technological change, Series F: Computer and systems sciences* (Vol. 141, pp. 112–124). Berlin, Germany: Springer.
- \*Frese, M., Brodbeck, F. C., Heinbokel, T., Mooser, C., Schleiffenbaum, E., & Thiemann, P. (1991). Errors in training computer skills: On the positive function of errors. *Human-Computer Interaction, 6*, 77–93.
- Frese, M., & Zapf, D. (1994). Action as the core of work psychology: A German approach. In H. C. Triandis, M. D. Dunette, & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 4, pp. 271–340). Palo Alto, CA: Consulting Psychologists Press.
- Goodman, J., & Wood, R. E. (2004). Feedback specificity, learning opportunities, and learning. *Journal of Applied Psychology, 89*, 809–821.
- \*Granados, A. R. (2000). *Fehlertraining in der Analyse: Die Wirkung von Heuristiken auf Motivation, Aufmerksamkeit und Leistung* [Analysis of error training: The effect of heuristics on motivation, attention, and performance]. Unpublished diploma thesis, Justus-Liebig University of Giessen, Germany.
- \*Greif, S., & Janikowski, A. (1987). Aktives Lernen durch systematische Fehlerexploration oder programmiertes Lernen durch Tutorials? [Active learning via systematic exploration of errors or programmed learning via tutorials?]. *Zeitschrift für Arbeits- und Organisationspsychologie, 31*, 94–99.
- Greif, S., & Keller, H. (1990). Innovation and the design of work and learning environments: The concept of exploration in human-computer interaction. In M. A. West & J. A. Farr (Eds.), *Innovation and creativity at work: Psychological and organizational strategies* (pp. 231–249). Oxford, England: Wiley.
- \*Gully, S. M., Payne, S. C., Koles, K. L. K., & Whiteman, J. A. K. (2002). The impact of error management training and individual differences on training outcomes: An attribute-treatment interaction perspective. *Journal of Applied Psychology, 87*, 143–155.
- Hacker, W. (1998). *Allgemeine Arbeitspsychologie: Psychische Regulation von Arbeitstätigkeiten* [General industrial psychology: Mental regulation of working activities]. Bern, Switzerland: Huber.
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*, 107–128.
- Hedges, L. V. (1982). Fitting categorical models to effect sizes from a series of experiments. *Journal of Educational Statistics, 7*, 119–137.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. London: Academic Press.
- \*Heimbeck, D., Frese, M., Sonnentag, S., & Keith, N. (2003). Integrating errors into the training process: The function of error management instructions and the role of goal orientation. *Personnel Psychology, 56*, 333–361.
- \*Heinbokel, T. (1990). *Das Konzept des Fehlermanagements in der Software-Schulung* [The concept of error management in software training]. Unpublished diploma thesis, University of Munich, Germany.
- Hesketh, B. (1997). Dilemmas in training for transfer and retention. *Applied Psychology: An International Review, 46*, 317–339.
- Hesketh, B., & Ivancic, K. (2002). Enhancing performance through training. In S. Sonnentag (Ed.), *Psychological management of individual performance* (pp. 249–265). New York: Wiley.
- Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.
- Ilgel, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology, 64*, 359–371.
- \*Irmer, C., Pfeffer, S., & Frese, M. (1991). Konsequenzen von Fehleranalysen für das Training: Das Fehlertraining [Consequences of error analysis for training: Error training]. In M. Frese & D. Zapf (Eds.), *Fehler bei der Arbeit mit dem Computer* (pp. 151–165). Bern, Switzerland: Huber.
- \*Ivancic, K. (1998). Errors as a means of promoting transfer of training (Doctoral dissertation, University of New South Wales, 1997). *Dissertation Abstracts International, 58*, 3832.
- Ivancic, K., & Hesketh, B. (1995/1996). Making the best of errors during training. *Training Research Journal, 1*, 103–125.
- Ivancic, B., & Hesketh, K. (2000). Learning from error in a driving simulation: Effects on driving skill and self-confidence. *Ergonomics, 43*, 1966–1984.
- Joung, W., Hesketh, B., & Neal, A. (2006). Using “war stories” to train for adaptive performance: Is it better to learn from error or success? *Applied Psychology: An International Review, 55*, 282–302.
- Kanfer, R., Ackerman, P. L., & Heggestad, E. D. (1996). Motivational skills and self-regulation for learning: A trait perspective. *Learning and Individual Differences, 8*, 185–209.
- \*Keith, N., & Frese, M. (2005). Self-regulation in error management training: Emotion control and metacognition as mediators of performance effects. *Journal of Applied Psychology, 90*, 677–691.
- \*Keith, N., & Müller, H. (2004, September). *Selbstregulation beim Fehlermanagement-Training am Computer: Emotionskontrolle als Mediator und Moderator der Trainingswirksamkeit* [Self-regulation in error management training on the computer: Emotion control as mediator and moderator of training effectiveness]. Poster session presented at the 44th Congress of the Deutschen Gesellschaft für Psychologie, Göttingen, Germany.
- Kirkpatrick, D. L. (1987). Evaluation of training. In R. L. Craig (Ed.), *Training and development handbook: A guide to human resource development* (3rd ed., pp. 301–319). New York: McGraw-Hill.
- Kluger, A. N., & DeNisi, A. (1996). Effects of feedback intervention on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*, 254–284.
- Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation [Monograph]. *Journal of Applied Psychology, 78*, 311–328.
- Latham, G. P., & Locke, E. A. (1991). Self-regulation through goal setting. *Organizational Behavior and Human Decision Processes, 50*, 212–247.
- \*Lazar, J., & Norcio, A. (2003). Training novice users in developing strategies for responding to errors when browsing the Web. *International Journal of Human-Computer Interaction, 15*, 361–377.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Lorenzet, S. J., Salas, E., & Tannenbaum, S. I. (2005). Benefiting from mistakes: The impact of guided errors on learning, performance, and self-efficacy. *Human Resource Development Quarterly, 16*, 301–322.

- Neubert, M. J. (1998). The value of feedback and goal setting over goal setting alone and potential moderators of this effect: A meta-analysis. *Human Performance, 11*, 321–335.
- \*Nordstrom, C. R., Wendland, D., & Williams, K. B. (1998). "To err is human": An examination of the effectiveness of error management training. *Journal of Business and Psychology, 12*, 269–282.
- Oswald, F. L., & McCloy, R. A. (2003). Meta-analysis and the art of the average. In K. R. Murphy (Ed.), *Validity generalization: A critical review* (pp. 311–338). Mahwah, NJ: Erlbaum.
- Rosenthal, R. (1991). *Meta-analysis procedures for social research* (Rev. ed.). Newbury Park, CA: Sage.
- Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology, 52*, 59–82.
- Salas, E., & Cannon-Bowers, J. A. (2001). The science of training: A decade of progress. *Annual Review of Psychology, 52*, 471–499.
- Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanisms of a neglected phenomenon. *Educational Psychologist, 24*, 113–142.
- Schmidt, F. L., & Hunter, J. E. (2002). Meta-analysis. In N. Anderson, D. S. Ones, et al. (Eds.), *Handbook of industrial, work, and organizational psychology: Vol. 1. Personnel psychology* (pp. 51–70). London: Sage.
- Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*, 207–217.
- Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning and Memory, 5*, 179–187.
- Skinner, B. F. (1953). *Science and human behavior*. New York: Free Press.
- Smith, E. M., Ford, J. K., & Kozlowski, S. W. J. (1997). Building adaptive expertise: Implications for training design strategies. In M. A. Quinones & A. Ehrenstein (Eds.), *Training for a rapidly changing workplace* (pp. 89–118). Washington DC: American Psychological Association.
- Sonnentag, S. (1998). Expertise in professional software design: A process study. *Journal of Applied Psychology, 83*, 703–715.
- \*Stiso, M. E., & Payne, S. C. (2007). *The influence of incentives and timing on the effectiveness of error training*. Unpublished manuscript.
- Taylor, P. J., Russ-Eft, D. F., & Chan, D. W. L. (2005). A meta-analytic review of behavior modeling training. *Journal of Applied Psychology, 90*, 692–709.
- \*Thiemann, P. (1990). *Aus Fehlern lernen? Fehlervermeidung vs. Fehlermanagement in der Mensch-Computer Interaktion* [Learning from errors? Error prevention vs. error management in human-computer interaction]. Unpublished diploma thesis, University of Munich, Germany.
- \*van Dyck, C. (2007a). *Error attribution in error management versus error aversive environments*. Unpublished manuscript.
- \*van Dyck, C. (2007b). *Mastery and aversion approaches to error disentangled from reinforcement effects: Confounds of the field translated to the lab*. Manuscript in preparation.
- van Dyck, C., Frese, M., Baer, M., & Sonnentag, S. (2005). Organizational error management culture and its impact on performance: A two-study replication. *Journal of Applied Psychology, 90*, 1228–1240.
- \*Wood, R. E., Kakebeeke, B. M., Debowski, S., & Frese, M. (2000). The impact of enactive exploration on intrinsic motivation, strategy, and performance in electronic search. *Applied Psychology: An International Review, 49*, 263–283.
- \*Yorke, C. M. (2006). *Toward minimizing negative consequences of errors: Exploring interactions between trainee characteristics and training design* (Doctoral dissertation, DePaul University, 2002). *Dissertation Abstracts International, 66*, 6324.

Received December 19, 2004

Revision received April 20, 2007

Accepted May 3, 2007 ■