

Error Orientation Questionnaire (EOQ): reliability, validity, and different language equivalence

**VOLKER RYBOWIAK¹, HARRY GARST², MICHAEL FRESE^{1*}
AND BERNAD BATINIC²**

¹*University of Amsterdam, The Netherlands*

²*University of Giessen, Germany*

Summary

An Error Orientation Questionnaire (EOQ) was developed, consisting of eight scales on attitudes to and on coping with errors at work. In Study I (representative sample of a German city, $N = 478$) six scales were developed with the help of a confirmatory factor analysis using LISREL techniques. They comprise error competence, learning from errors, error risk taking, error strain, error anticipation, covering up errors. All constructs were validated. In a second study, items were added to the scales and two additional scales, 'error communication' and 'thinking about errors', were included. The scales were translated into English and Dutch and 160 students in the Netherlands filled out both language versions (Study II). The 8-factor solutions in English and Dutch were replicated. The issue of language equivalence of these two language versions were taken up (equivalence across correlations exists). Potentially biasing variables did not influence the solutions. Practical uses of the EOQ are pointed out. Copyright © 1999 John Wiley & Sons, Ltd.

Introduction

It is our objective to present the development of scales that measure how one copes with and how one thinks about errors at work. Since we believe that this instrument should be used in various European countries, we thought it would be necessary to present it in different languages from the start. Therefore, we will present the construction of reliable and comparable scales in

* Correspondence to: Michael Frese, Department of Psychology, Roetersstraat 15, 1018 WB Amsterdam, Netherlands.
Fax: 0641 71166. E-mail: volker.rybowiek@psychol.uni-giessen.de

The EOQ is available in Dutch, English and German from Michael Frese.

Contract Grant Sponsor: European Human Capital Mobility Grant (Prevention of Human Error in Systems for Energy Production and Process Industry).

Contract Grant Number: ERBCHRXCT 930298.

Contract Grant Sponsor: Deutsche Forschungsgemeinschaft (DFG).

Contract Grant Number: Fr 638/6-5.

German, Dutch and English—the Error Orientation Questionnaire (EOQ). In addition, the constructs have been validated.

Errors are an important issue in work psychology for various reasons. First, it is the raw material that produces stress, accidents, inefficient human–machine interaction, quality and performance problems, and a bad climate. Thus, many recurrent problems in industry are related to the issue of errors, for example, the Three Mile Island accident or Leeson's 1.5 billion dollar loss. Second, attitudes towards errors and how one deals with them are indications of a company's organizational culture. Bureaucratic companies usually attempt to prevent errors at all costs, while entrepreneurial cultures have a more positive attitude towards errors and what one can learn from them. Thirdly, if a company attempts to change its culture or if one wants to introduce the issue of error in selection procedures, one needs a measure of error orientation.

Whether one perceives errors as negative indicators of performance or not is of major importance. If errors are not allowed to appear in a company, one has to develop constant checks (technical systems or other people), there will be little risk taking, and there will be a high degree of planning and little action (Peters, 1987). The major weakness of such an approach is that errors cannot be eradicated (cf. Frese and Zapf, 1991). Neither routinization nor qualification protect a company from errors (Prümper, Zapf, Brodbeck and Frese, 1992). In certain negative error cultures, those errors that appear will be concealed and there will be little individual and organizational learning from errors. Sometimes, error prevention systems may actually increase the negative effects of errors, because people do not expect errors to occur (Bainbridge, 1983; Reason, 1990). In contrast, if a company has a more positive attitude towards errors, it can be more action-oriented, innovative, and experimental.

Other important issues for the company are, for example, whether or not errors are anticipated (error anticipation is related to better handling of errors, Grefe, 1994) or the number of risks people take when making errors. Error orientation may very well turn out to be one of the most important organizational background variables, because the way an organization deals with errors may determine the amount of organizational learning, since, for example, errors used as enhancers of learning lead to better performance (Dormann and Frese, 1994) and, therefore, to greater organizational effectiveness (Argyris, 1990). For organizational learning from errors, it is necessary to use active approaches to errors rather than more passive approaches (Frese, 1995). A learning organization (Senge, 1990) needs to have a positive attitude towards exploration and errors and to deal with errors actively. Moreover, an organization's error orientation may affect the possibility of introducing modern quality assurance programs (e.g. quality circles, or Kaizen). The degree to which workers are actually willing to take on added responsibilities that follow from modern production techniques (e.g. lean production: Womack, Jones, Roos and Sarpenter, 1990) is determined by whether or not errors lead to negative consequences for the agent. On a more macro level, Hofstede (1991) has speculated that uncertainty avoidance is an important variable that distinguishes, for example, Germany from other countries within Western Europe. The amount of anxiousness produced by potential errors may well be a reason for uncertainty avoidance.

Thus, error orientation is an interesting variable, on an individual level, a company level, and even a cultural level. To our knowledge there are no scales or questionnaires developed to measure orientation towards errors. Those scales on errors that do exist refer to a person's general absentmindedness, for example, the Cognitive Failure Questionnaire (CFQ by Broadbent, Cooper, Fitzgerald and Parkes, 1982). Therefore, they cannot be used to study error orientation.

Error orientation can be conceptualized within a general coping concept (Lazarus and Folkman, 1984). Primary appraisal is related to how negatively errors are perceived and the

degree to which one anticipates that errors will happen. Secondary appraisal refers to coping with errors. Coping strategies are to calm oneself in the face of errors, to cover up the fact that an error has occurred, as opposed to communicating about them, and to actively deal with an error or to learn from it. Thus, the primary and secondary appraisal concepts helped us to develop items even though we did not yet have a fully-developed theoretical framework. However, in an empirical exploratory study on errors by managers, the aforementioned categories have been shown to exist, at least in the descriptions that managers gave of their errors (Grefe, 1994).¹

Thus, it is the goal of this article to establish a differentiated questionnaire to study error orientation with a set of factors. The error orientation questionnaire is intended to be useful for both practical and theoretical purposes.

In terms of construct validation, the error orientation constructs should have meaningful correlations with coping resources and with strain. Errors are stressful (Brodbeck, Zapf, Prümper and Frese, 1993), one reason being that they disrupt goal-directed behavior (Frese and Zapf, 1994; Leitner, Lüders, Greiner, Weber and Hennes, 1987; Semmer, 1984). Error orientation may be related to depression and strain in general. In contrast, active error-coping strategies should be related to a general pro-active orientation towards working life. However, there are also differences between typical stressors at work and errors. First, errors are self-produced by the agent and, therefore, there is more blame on oneself. Second, the negative error consequences are often made public. Finally, errors often disrupt an action in such a way that one has to deal with the error. Thus, pure emotion-oriented coping strategies have clear limits—the problem has to be solved if the goal is still to be pursued.

Coping resources may lie in the person or in the work situation. Personal resources may be self-efficacy and self-esteem, good planning skills, readiness to change at work, and high action orientation. A work-related personal resource is qualification. On the other hand, lack of organizational resources may be related particularly to covering up errors. For example, if one's career is jeopardized or if one is liable to lose one's job, there may be little tendency to communicate one's errors at the workplace. Thus, we assume that coping resources should be related meaningfully with the error orientation constructs. For reasons of exposition, the exact hypotheses on construct validation will be described after the scales have been developed as a result of Study I, because only at this point do we know the exact factorial structure of the EOQ (cf. Results on Construct Validation).

To pursue our goal of developing the EOQ and to look into the construct validity and reliability, two studies will be reported on.

Study I

Methods

Sample

This study used the sample of the longitudinal project AHUS (German acronym of 'active actions in radical change situations') at time 5. (NB. the main objective of the AHUS project is the investigation of the developing of active strategies and personal initiative with changing working conditions in East Germany, cf. Frese, Kring, Soose and Zempel, 1996). The representative

¹ We would like to thank Judith Grefe for drawing up the first version of the item set.

sample was taken in Dresden (East Germany) by randomly selecting streets, taking every third house and in each house every fourth apartment, to gain subjects 18- to 65-year-olds with full-time employment (at time 1). The refusal rate (33 per cent) was relatively low. At t1 463 subjects and at t2 202 additional people took part; at t5 the N was still 478. For reasons of LISREL-analytical cross-validation, the sample was divided into two subsamples of $N = 239$ each.

Procedure

Thirty-two items covering cognitive, affective and behavioral aspects of attitudes and coping with regard to errors were given (the full scales are listed in the Appendix). To establish a reasonable measurement model, we used a comprehensive strategy, suggested by Gerbing and Hamilton (1996), that exploits the relative and complementary strength of exploratory and confirmatory factor analytic procedures for model construction, evaluation, and revision. They have shown that their approach was superior to using only confirmatory factor analysis. In our study three steps were taken: first, a classical exploratory factor analysis was carried out to provide a preliminary classification. This was necessary to understand the underlying structure of the indicators. Second, we used LISREL techniques to fine tune the model. Since we modified our scales empirically, this represented an exploratory strategy and it was, therefore, restricted to the calibration sample. Third, the models were then cross-validated in the validation sample, as suggested by Cudeck and Browne (1983).

Construct validity was established by developing a nomological net for each construct (Cronbach and Meehl, 1955), i.e. the EOQ scales were correlated with constructs to which they should have theoretically plausible relationships.

Measures

Most scales used for construct validation were based on a questionnaire. *Self-efficacy* was developed as work-related generalized self-efficacy (Speier and Frese, 1996). *Self-esteem* by Mohr (1986) is a highly generalized work-related self-confidence scale. *Plan-orientation* describes the tendency to plan actions in the long run and in detail (Frese, Stewart and Hannover, 1987). *Action-orientation after failure* means quick actions as a consequence of intentions one sets up after failure (Kuhl, 1983).

Readiness for change implies that one is willing to accept or to look for changes at work (Frese and Plüddemann, 1993). *Control rejection* by Frese (1984) was used as the reverse of taking responsibility. It means that one does not want more autonomy at work because this would increase one's responsibilities (Frese, Erbe-Heinbokel, Grefe, Rybowskiak and Weike, 1994). *Need for achievement* is a measure for the consequent pursuit and realization of high goals at work (Modick, 1978). *Career stress* is a measure for difficulties and obstructions in one's career. *Job uncertainty* is an assessment of how likely one is to lose one's job. Both scales were developed for this study. *Qualification* is an overall estimate that the interviewers made after exploration of the subjects' careers from the beginning of their working life to their present job.

Psychosomatic complaints is a revised version of Fahrenberg's scale (1975) by Mohr (1986) and comprises, for example, tensions and cardiac disorders. *Depression* (Mohr, 1986) refers to sad moods and depressing thoughts. *Negative affectivity*, measured by a short form of the PANAS with the generalized instruction (Watson, Clark and Tellegan, 1988), describes mental conditions resulting from a generalized neuroticism, such as feelings of shame, guilt, fear etc. *Optimism* (as a translation of the Life Orientation Test, Scheier and Carver, 1985) means that one has general positive expectancies and hopes.

Results on the development of the EOQ scales

First, a preliminary classification was made with the help of an exploratory principal component analysis. This leads to eight factors (NB. the loadings had to be greater than 0.40 and to differ at least 0.10 to the loading on another factor). These eight factors were then analyzed with a confirmatory factor analysis (LISREL). After testing the eight models separately the constructs were then combined in pairs (Jöreskog and Sörbom, 1993). The following fit measures were used: chi-square, goodness-of-fit-index, adjusted-goodness-of-fit-index, root-mean-square-residual and expected-cross-validation-index, in addition, to standard errors, *t*-values and modification indices. At each step, modifications such as elimination of indicators or changing paths from latent variables to indicators were carried out.

Four alternative and theoretically plausible models were developed that describe all constructs jointly. In the last analysis we had to decide between two best fitting models, one with eight and one with six factors (model B with eight and model C with six factors). Model C was chosen because it had the better fit measures. It did not include the constructs of error communication and thinking about errors, which both had only two indicators in Study I. Model C has good fit indices, given the complexity of such a model. (It also has a good expected-cross-validation-index, ECVI; Browne and Cudeck, 1993, cf. Table 1 calibration model).

Table 1. Fit measures of the calibration model and the cross-validation models, Study I

	Chi ²	df	<i>p</i> -value	GFI	AGFI	stand. RMR	ECVI
Calibration model							
C	180.49	135	0.0054	0.94	0.91	0.044	1.06
Cross-validation models							
C	531.01	325	—	0.89	—	0.074	1.17
C*	435.19	302	0.0000077	0.92	—	0.060	1.08

GFI, Goodness of Fit Index; AGFI, Adjusted Goodness of Fit Index; stand. RMR, Standardized Root Mean Square Residual; ECVI, Expected Cross Validation Index.

However, since model C was generated with an exploratory strategy, it needs to be cross-validated. The cross-validated model C used the same constraints as the calibration model (cf. Table 1). A legitimate improvement in the fit indices can be achieved by setting the error variances free and allowing a few extra paths in the validation sample (in bold type in Figure 1). This was done in model C*, presented in Table 1 and in Figure 1. One can see that there are adequate fit measures in this cross-validated model,² again given the complexity of the full model and that the extra paths are of negligible magnitude. Thus, we were able to cross-validate the factor structure across these two samples.

Results on the construct validation of the EOQ scales

Our measurement model of the EOQ contains six constructs to be validated by establishing theoretically driven nomological nets (Cronbach and Meehl, 1955). The relevant expected

²The ratio of the value of chi² and degrees of freedom is slightly above 1, which can be regarded as a good fit, even if the chi² is significant.

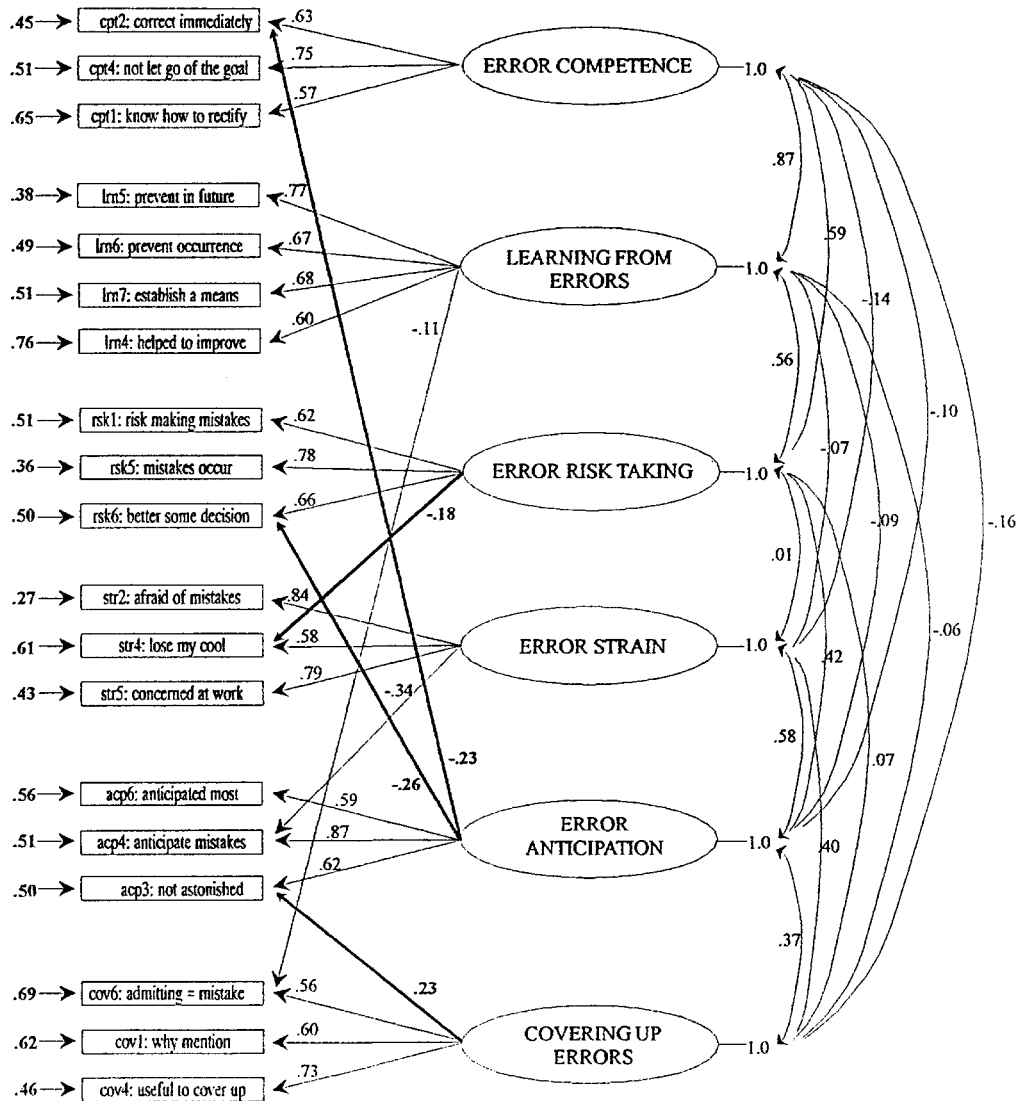


Figure 1. Measurement model C* tested (on invariance over groups) on the validation sample, Study I. The resulting differences due to modification of model C are shown in bold type

correlations for these nomological nets are discussed here for reasons of exposition, since the scales have not been introduced prior to this section. All relevant correlations are displayed in Table 2.

Error competence and learning from errors

The distinction between the two constructs should be as follows: error competence means the knowledge of and capability to deal with errors *immediately*, when they occur. Thus, it is directed at short-term goals. Therefore, there should be a positive correlation between error competence and action-orientation after failure (Kuhl, 1983). In contrast, ‘learning from errors’ refers

Table 2. Correlations of the EOQ scales with other meaningful variables, Study I

	Error competence	Learning from errors	Error risk taking	Error strain	Error anticipation	Self-efficacy	Plan orientation	Action orientation after failure	Qualification	<i>M</i>	<i>S.D.</i>
Error competence						0.44*	0.19*	0.20*	0.06	4.07	0.49
Learning from errors	0.63*					0.37*	0.24*	0.10‡	0.12†	4.11	0.53
Error risk taking	0.43*	0.44*				0.35*	0.13†	0.14†	0.24*	3.84	0.73
Error strain	-0.12†	-0.05	-0.04			-0.36*	-0.01	-0.46*	0.03	2.35	0.72
Error anticipation	-0.10‡	-0.05	0.28*	0.33*		-0.11†	0.03	-0.10‡	0.06	2.23	0.65
Covering up errors	-0.14†	-0.09‡	0.03	0.31*	0.26*	-0.16*	-0.08‡	-0.17*	0.13†	2.00	0.71

	Readiness for change	Control rejection	Need for achievement	Psychosomatic complaints	Depression	Negative affectivity	Optimism	Self-esteem	Job uncertainty	Career stress
Error competence	0.24*	-0.28*	0.46*	0.00	-0.27*	-0.14†	0.25*	0.42*	-0.17†	0.04
Learning from errors	0.28*	-0.31*	0.45*	0.01	-0.24*	-0.14†	0.23*	0.33*	-0.03	0.08
Error risk taking	0.29*	-0.22*	0.32*	-0.04	-0.19*	-0.09‡	0.14†	0.17*	-0.06	-0.02
Error strain	-0.22*	0.42*	-0.07	0.32*	0.44*	0.47*	-0.27*	-0.40*	0.18*	0.09
Error anticipation	-0.07	0.20*	-0.09‡	0.09‡	0.23*	0.22*	-0.23*	-0.26*	0.05†	0.12‡
Covering up errors	-0.09‡	0.30*	-0.14†	0.08‡	0.26*	0.23*	-0.20*	-0.24*	0.17†	0.22*

* $p < 0.001$; † $p < 0.01$; ‡ $p < 0.05$.

to learning something for the future, so that the work plans are optimized *in the long term*. Consequently, a positive correlation of learning from errors should appear with plan-orientation (Frese *et al.*, 1987). This is the tendency to plan actions for a longer period and in great detail. Both constructs deal with competence; therefore, they should be correlated positively with self-efficacy (i.e. the expectancy of one's competence; Bandura, 1986), with self-esteem and with qualification (i.e. competence related to one's job).

Results confirm our expectations to a certain extent (cf. Table 2). Both scales correlated with self-efficacy (0.44, 0.37), with self-esteem (0.42, 0.33), with plan-orientation (0.19, 0.24) and action orientation after failure (0.20, 0.10). The correlations with qualification were minimal, however (0.06, 0.12). This may be so because qualification was a very rough overall estimate of the qualification of the person (highly correlated with education). The differential expectation of the correlations with plan-orientation and action-orientation after failure did not prove to be correct. While error competence was correlated somewhat more closely than learning from errors with action-orientation after failure (0.20 versus 0.10) (difference in correlations $z = 1.52$, $p < 0.10$), the other correlations did not confirm our expectations. There were also high correlations with need for achievement. This makes sense because error situations present challenges and therefore, trigger the achievement motive (McClelland, 1987). Error competence and learning from errors were highly correlated ($r = 0.63$, cf. Table 2). Therefore, the differentiation between both constructs may be quite subtle.

Error risk taking

There are two theoretical expectations that may complement each other. (a) Error risk taking implies general flexibility and openness towards errors and should, therefore, correlate positively with readiness for change (Frese and Plüddemann, 1993) and negatively with control rejection (Frese *et al.*, 1994). This should be so because one is willing to adjust to new conditions at work and to take responsibility despite potential negative consequences. Hypothesis b suggests that error risk taking is due to high achievement-orientation, which implies that one accepts errors and their consequences in order to reach higher goals. Therefore, positive correlations with need for achievement are to be expected. The results support both expectations (cf. Table 2); there were correlations with readiness for change (0.29), control rejection (-0.22), and need for achievement (0.32). The correlations with self-efficacy (0.35) and qualifications (0.24) show that there is a certain degree of self-assuredness and competence in people who take error risks.

Error strain

Error strain means being strained by making errors and therefore fearing the occurrence of errors or reacting to errors with high emotions. As a result, there should be positive correlations with other strain measures, such as psychosomatic complaints, depression, and negative affectivity. The results in Table 2 clearly support these expectations; psychosomatic complaints ($r = 0.32$), depression (0.44), and negative affectivity (0.47) were all related to error strain. The negative correlation with self-esteem and self-efficacy also indicates the same picture: error strain is not highly specific to errors but related to more general strain measures.

Error anticipation

There are two competing perspectives with regard to error anticipation. (a) A general expectancy that errors will happen, because one has a realistic view that even in one's field of expertise errors

will occur. Given this perspective, there should be positive correlations with error competence and learning from errors. (b) A stable negative attitude to errors as a result of general pessimism, lack of self-esteem, and negative affectivity.

The results in Table 2 show perspective b to be true; error anticipation was not significantly related to error competence and learning from errors, but significantly correlated with optimism (-0.23), self-esteem (-0.26), and negative affectivity (0.22).

Covering up errors

There are two potentially complementary interpretations of the construct covering up errors. (a) A strategy of anxious people who consider errors as a threat and, therefore, prevent accusations often associated with errors in a social context (Hamilton, 1980). Consequently, there should be a negative correlation with self-esteem and a positive correlation with control rejection, because an insecure person would not be willing to take responsibility for her or his errors for fear of the consequences. (b) It could also be a reaction to certain organizational conditions. For example, it may be useful not to admit one's errors, if a workplace is at risk or when one's career is threatened. This implies positive correlations with job uncertainty and with career stress.

Results suggest that both interpretations are correct (cf. Table 2). Covering up was related to self-esteem (-0.24), to control rejection (0.30), to career stress (0.22), and to job uncertainty (0.17).

Discussion of Study I

We were able to develop a set of six scales—error competence, learning from errors, error risk taking, error strain, error anticipation, covering up errors—for measuring orientation towards errors. They were empirically developed but a cross-validation of the factor structure worked well. A first validation suggests that these scales behave in a law-governed manner. Thus, error competence and learning from errors were related to self-efficacy, plan- and action-orientation; error risk taking was related to need for achievement, control rejection, and readiness for change; error strain was related to ill-health; error anticipation implied a negative outlook on life; and covering up was related both to environmental factors (career stress, job insecurity) and to lack of self-esteem, rejecting responsibility and control.

The strength of this study is that we could cross-validate the factor structure and that the scales were developed in a representative sample. The limitation was that the scales were based on only three items, and two constructs (error communication and thinking about errors), which are theoretically important, could not be developed because there were not enough items in them. Therefore, an extended version was developed. Furthermore, for cross-cultural purposes for use in Europe, we also wanted to have English and Dutch versions of the scales. For this reason an extended German version with eight scales, including error communication and thinking about errors was translated into English and Dutch.

Study II

Study II had the following objectives: first, it should look at the English and Dutch versions of the EOQ and the psychometric properties of the scales. Second, we wanted to know whether an

equivalence of the two language versions could be achieved. Third, the EOQ was to be related to initiative at work.

Since we wanted to study both language versions at the same time, we needed a sample that was bilingual. Dutch university students are nearly bilingual. Some of their courses are in English, nearly all of the reading material is in English,³ and much of the television program is in English. In general, when one approaches the students in English, they can nearly always give an English answer immediately and without hesitation. Thus, we thought that Dutch students could be used to study the equivalence of the factorial structure across two languages. This design allows us to separate language equivalence issues from cultural equivalence. Usually in cross-cultural research, one is worried about language and cultural equivalence issues that may explain differences in results (Poortinga, 1989). In this study, we only have one culture but two languages. Thus, differences are due to language.

In addition, we used a subjective measure of initiative (Frese, Fay, Hilburger, Leng and Tag, 1997). Since initiative means that one actively changes things, it implies that one will make more errors. Only if one is not fearful of and stressed by errors and confident that one can handle them, will one show initiative. Thus, high initiative people should be less strained by errors (negative relationship with error strain), show higher error competence and learning from errors. Additionally, they should be more willing to talk about their mistakes and risk mistakes (negative relationship to covering up and positive correlation with error risk taking).

Method

Translation

The German version was translated into English by the authors with the help of native English speakers. This was then translated back into German and a committee of three researchers changed the English version as a result. (The English version is given in the Appendix). The Dutch version was translated from the English and the German versions independently. The initiative scale was translated from German into Dutch.

Sample and design

The sample consisted of 160 subjects, 76 males and 84 females, 13 econometrics students, 59 polytechnic part-time students, and 34 other subjects (usually acquaintances of the students who were collecting the data⁴). The polytechnic students were rewarded with lottery tickets and the psychology students had to fulfill the requirement to participate in psychological research as a subject. The age of the subjects ranged from 19 to 55 years with an average age of 26.5. The majority (74 per cent) of the subjects had (part-time) jobs.

To prevent order effects, the subjects were randomly assigned to two groups. The two language versions were presented in opposite succession to both groups. In order to prevent artificial dependence among the items of both languages, the subjects were asked during the session not to look back to previous results. For the same reason, the initiative scale was presented between the two language versions of the EOQ.

³There are differences between universities and polytechnic (HES, Hogere Economische School); English language material is used much less in the HES.

⁴A bias analysis along the lines discussed later in this article, comparing these 34 subjects with the rest, showed that these 34 subjects did not affect the factor structure. Thus, these 34 people are comparable to the other subjects.

Results

We checked for outliers but there were none. We also checked for the distributions of the items but there was no threat of high kurtosis. The same goodness of fit measures as in Study I were used.

We shall report results gained on five issues: the first part is concerned with the psychometric properties of the individual Dutch and English scales. The second is related to the equivalence of the two language versions. A third part discusses the discriminant validity of the scale items. (Do the items measure the intended constructs and not the others?) The fourth part reports on potential nuisance factors influencing our results, for example, influence of language ability, time of work experience, age, sex, and order effects on item bias. Finally, we report correlations of the EOQ scales with initiative.

Psychometric properties of scales

The psychometric properties of the individual Dutch and English scales are reported in Table 3. The alphas were nearly all above 0.70 and, therefore, adequate for research purposes. One exception was error competence, which was only 0.56 in the English version. In most cases, the English language scales have lower alphas than the Dutch ones. This is understandable since English was not our subjects' native language (although we maintain that Dutch university students are the nearest to bilingual subjects one can get). For this reason, the comparatively low reliability of the English error competence scale may be acceptable. The results look promising for the English version, but the quality of the scales still has to be established with native speakers.

For each scale in each language a confirmatory factor analysis was performed separately to test whether a 1-factor model fitted properly. The goodness of fit measures were generally acceptable. Error competence (Dutch and English) and error risk taking (English version) were an exception. The chi-square-to-*df*-ratio of these scales was relatively high, their AGFI are the lowest, and their RMR the highest (cf. Table 3). The modification indices revealed that there were intercorrelations of the error terms between two items of the error competence scale. This is so because these two items are phrased rather similarly, producing a common factor between these two items. However,

Table 3. Means, standard deviations, alphas and goodness of fit measures of the scales, Study II

	<i>M</i>	<i>S.D.</i>	Alpha	Goodness of fit						
				Chi ²	<i>df</i>	<i>p</i>	GFI	AGFI	RMR	CFI
Error competence Dutch	3.72	0.54	0.71	10.22	2	0.006	0.97	0.85	0.055	0.93
Error competence English	3.59	0.53	0.56	12.89	2	0.0016	0.96	0.82	0.072	0.83
Learning from errors Dutch	3.51	0.68	0.78	6.18	2	0.045	0.98	0.90	0.034	0.97
Learning from errors English	3.53	0.77	0.89	0.11	2	0.95	1.00	1.00	0.003	1.00
Error risk taking Dutch	3.63	0.75	0.82	1.37	2	0.50	1.00	0.98	0.014	1.00
Error risk taking English	3.48	0.72	0.74	14.39	2	0.000	0.96	0.79	0.057	0.91
Error strain Dutch	2.48	0.85	0.86	16.63	5	0.0053	0.96	0.88	0.038	0.97
Error strain English	2.51	0.79	0.79	8.84	5	0.12	0.98	0.93	0.037	0.98
Error anticipation Dutch	3.19	0.75	0.82	11.05	5	0.05	0.97	0.91	0.036	0.97
Error anticipation English	3.02	0.68	0.73	11.59	5	0.041	0.97	0.92	0.044	0.97
Covering up errors Dutch	2.19	0.70	0.83	15.84	9	0.07	0.97	0.92	0.039	0.98
Covering up errors English	2.27	0.69	0.78	23.78	9	0.0047	0.95	0.89	0.053	0.93
Error communication Dutch	3.52	0.66	0.71	4.09	2	0.13	0.99	0.94	0.034	0.98
Error communication English	3.41	0.67	0.67	5.53	2	0.063	0.98	0.92	0.036	0.96
Thinking about errors Dutch	3.56	0.65	0.84	5.04	5	0.41	0.99	0.96	0.023	1.00
Thinking about errors English	3.37	0.70	0.83	15.44	5	0.009	0.96	0.89	0.046	0.96

both items are reliable and we, therefore, kept them and this scale intact. With respect to the English version of the error risk taking scale, we could not identify any cause for the impairment of fit. Nevertheless, given some satisfying indices (GFI, CFI) and the fact that the comparative fit of alternative models is more important than focussing on the absolute fit of any single model (Tanaka, 1993), the scale is still acceptable, though it needs some further consideration.

Equivalence of the two language versions

Up to this point, each language version was looked at separately and it was shown that the EOQ scales can be reliably used both in Dutch and in English. We are now interested in language equivalence, which would allow cross-cultural comparisons. There are three equivalent issues: (1) mean differences, (2) item by item equivalence, and (3) scale equivalence.

Mean equivalence We created eight difference variables by subtracting the added scale scores of the English from the Dutch scales. We then performed a Multivariate Matched Pairs Analyses (Stevens, 1992)—a multivariate analysis of variance (MANOVA) that tests whether the joint effects are different from zero. The MANOVA showed a significant result ($F(146,8) = 9.22$, $p < 0.001$). A *post hoc* procedure consisting of paired *t*-tests showed that for most scales, the Dutch versions had higher mean scores than the English versions. (The Dutch and English means, *S.D.*s, and *t*-test results are presented in Table 4). Exceptions were learning from errors and error strain, which showed no significant differences, and covering up errors, which had a significantly higher score in the English version. In addition to the scales, *t*-tests for differences on the *item* level were done and several items differed significantly. An inspection of the translation showed that for those items there was a difference in the phrasing of the Dutch and the English items. Dutch phrasing was closer to the original German, whereas in the English version, the translation was somewhat less literal. For example, the word-by-word translation of the German item *cpt1* would have been 'When I have made a mistake, it mostly occurs to me how to make it good again'. The actual English translation is 'When I have made a mistake, I know immediately how to correct it'. Given the same degree of error competence, the statement in the English formulation may force less agreement on the scale because the wording is a bit 'stronger' here. In this respect the Dutch translation resembles the German version more.⁵

Table 4. Differences in scale scores between Dutch and English versions, Study II

Scales	Mean of Dutch scale	Mean of English scale	<i>S.D.</i> of difference score	<i>S.E.</i> of difference score	<i>t</i>	<i>df</i>	<i>p</i>
Error competence	3.72	3.59	0.42	0.034	4.1	156	0.000
Learning from errors	3.51	3.53	0.56	0.045	-0.6	156	0.522
Error risk taking	3.63	3.48	0.43	0.034	4.3	153	0.000
Error strain	2.48	2.51	0.39	0.031	-1.2	156	0.235
Error anticipation	3.19	3.02	0.50	0.040	4.0	154	0.000
Covering up errors	2.19	2.27	0.46	0.037	-2.4	155	0.016
Error communication	3.52	3.41	0.52	0.041	2.5	155	0.013
Thinking about errors	3.56	3.37	0.43	0.035	5.4	156	0.000

⁵In a study by Bennett (1977) bilinguals also tended to give lower scale scores in English than in their native language. However, Bennett speculated that the differences relate to a switching of reference groups depending on the questionnaire's language of presentation.

Table 5. Goodness of fit measures of the combined models, Study II

	<i>f</i>	Goodness of fit							Dchi ²	Ddf
		Chi ²	<i>df</i>	<i>p</i>	GFI	AGFI	RMR	CFI		
Error competence										
Model 1	0.89	48.79	15	0.000	0.93	0.84	0.076	0.90		
Model 2	0.87	50.97	19	0.000	0.93	0.87	0.08	0.91	2.18	4
Learning from errors										
Model 1	0.83	18.69	15	0.23	0.97	0.93	0.031	0.99		
Model 2	0.82	31.47	19	0.036	0.95	0.91	0.093	0.98	12.78*	4
Error risk taking										
Model 1	1.00†	30.17	16	0.017	0.96	0.90	0.038	0.97		
Model 2	1.00†	42.23	20	0.003	0.94	0.89	0.08	0.96	12.06*	4
Error strain										
Model 1	0.98	47.65	29	0.016	0.94	0.89	0.045	0.98		
Model 2	0.98	68.93	34	0.000	0.92	0.86	0.092	0.96	21.28*	5
Error anticipation										
Model 1	0.92	73.61	29	0.000	0.91	0.83	0.061	0.92		
Model 2	0.92	85.35	34	0.000	0.90	0.83	0.089	0.91	11.74*	5
Covering up errors										
Model 1	0.92	87.51	47	0.000	0.92	0.86	0.056	0.94		
Model 2	0.93	90.41	53	0.000	0.91	0.87	0.061	0.95	2.90	6
Communication about errors										
Model 1	0.82	15.62	15	0.41	0.98	0.94	0.036	1.00		
Model 2	0.82	16.24	19	0.64	0.98	0.95	0.039	1.00	0.62	4
Thinking about errors										
Model 1	0.92	42.65	29	0.049	0.95	0.91	0.039	0.98		
Model 2	0.92	44.12	34	0.11	0.95	0.92	0.049	0.99	1.47	5

* $p < 0.05$.

† Phi parameter fixed to the value of 1 to prevent an improper solution.

Item-by-item equivalence Figure 2 describes an example for the confirmatory factor analyses (LISREL) that were done for each scale to look at the English and Dutch versions' equivalence. The residual factors were left free to correlate with the other language item (a covariance matrix was used). This was done because the two language items share a common factor on the item level (Bollen, 1989), for example, the item 'When I do something wrong at work, I correct it immediately', has a common meaning in both languages.

The results of these analyses are reported in Table 5. Model 1 implies that all parameters were freed. Model 2 restricted the factor loadings of the items across the two languages so that they were the same.

The results of the model 1 of each scale show that the fit indices are good for learning from errors, error risk taking, error strain, error communication and thinking about errors and partly acceptable for error competence. The fit indices of the models measuring covering up errors and the anticipation of errors were lower than the fit indices for the separate models (cf. Table 3). Here the modification indices suggested that there were correlations of the error terms. These correlations were to be expected between the same items in the different languages (as shown in Figure 2). However, there were also correlations between unrelated items across the two

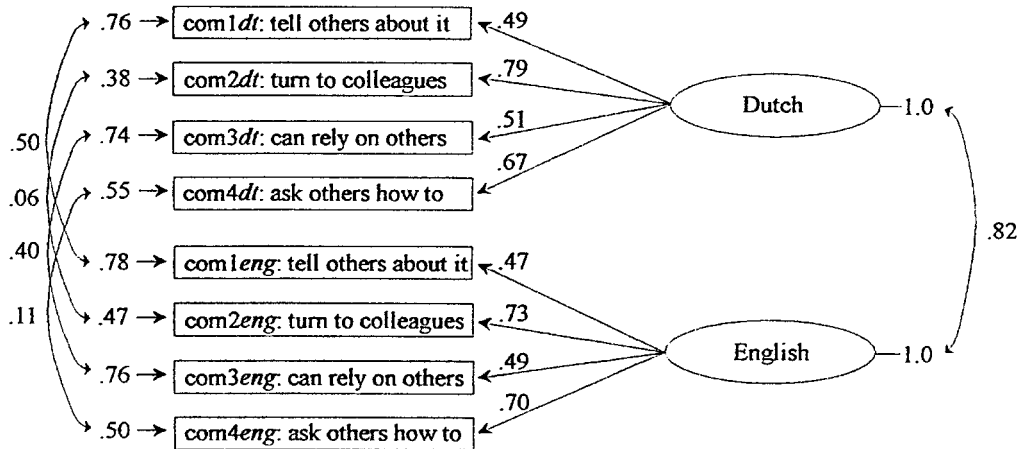


Figure 2. Combined measurement model for the Dutch and English error communication scale, Study II

languages (e.g. between item cov2 in the Dutch version and cov5 in the English version of covering up errors). Since this is a methodological problem that only appears when presenting the scales in different language versions to the same subjects, we are not so concerned about it. Normally, one will only use either the English or the Dutch scales, and then error terms between these two languages will not appear.

The results of model 2 in Table 5 can be tested against model 1 because they are nested (chi-square difference test, Bollen, 1989). The results indicate that model 2 is significantly worse than model 1 in four of the eight cases. Thus, there is little equivalence on an item by item basis if the restriction of equal loadings is introduced. However, model 1, which is less restrictive, is adequate because one would not necessarily suppose that the factor loadings should be the same across different languages. It is much more important that the scales are equivalent than that the item loadings are equivalent.

Scale equivalence The correlations between the same latent constructs in different languages can be interpreted as scale equivalence coefficients (cf. the first column in Table 5). All of the correlations between the latent constructs were above 0.80, with most being above 0.90. This suggests that, on the level of the scale, there is a high degree of equivalence.

The overall results of the equivalence issue present a complicated but well interpretable picture: the English items have a lower mean than the Dutch items. Thus, mean comparisons between cultures are difficult to make with the EOQ at this stage in the development. (NB. only our design allows one to make this kind of decision, because we have been able to separate the language from the cultural issue.) On the other hand, we think that on a correlational level, cross-cultural comparisons can be made with the EOQ. The main reason is that correlations between the constructs of both language versions are quite high (Table 5). One could argue against our viewpoint by referring to the fact that the items do not have equivalent factor loadings (or error variances). Two arguments speak against this claim. First, we usually do not use factor scores but unweighted sums of items when reporting relationships with other variables. Thus, differences in factor scores are not of particular importance if the differences are not high. Second, Schwarzer, Bässler, Kwiatek, Schröder and Zhang (1997, p. 81) have argued in a similar situation that there should not be a demand for an item-by-item equivalence across languages because the

translators are not usually asked to make literal translations but should ‘understand the theoretical construct . . . and to find meaningful adaptations instead of literal translations’.

The discriminant validity of the scale items

To understand whether items were loading on the wrong latent construct, confirmatory factor analyses were used in the following way: the items of one scale were included in one analysis with all the other constructs. (This was done because the ratio of subjects to the free parameters would have been too unfavorable for a complete model including all the items in this study.) The modification indices help to see whether a path is wrongly specified. These analyses were done consecutively for all the scales.

There were 11 significant modification indices. The estimated factor loadings were not higher than 0.35. There was only one exception to this result, namely item rsk5 of the risk scale loaded high on the anticipation scale. Considering the content of the item, this was reasonable and we, therefore, transferred this item to the anticipation scale (item acp2). As a matter of fact, in the results presented above, this has already been done (but for purposes of demonstration, this result is presented here).

The results on the discriminant validity of the items are good and typical for analyses that look at scales from the same content domain (the domain of errors in our case).

Item bias produced by influence of language ability, work experience, age, sex and order effects

In order to have a reliable set of scales, one should test that the scores are not affected by constructs which theoretically should have no influence on error orientation. To test item bias, we looked at the direct influence of language ability, work experience, age, sex and order effects on the item functioning of each of the scales. The potentially biasing variables were included in the model as latent constructs measured without measurement error. The correlation between the trait and the potentially biasing variables was set free, because in practice there can be a relation between them. The presence of a direct path from the biasing variables to the item is an indication that the item is biased. This effect implies that a higher item score is not necessarily related to a higher value of the trait (Oort, 1996). There was no influence of language ability, work experience, and age on any of the items. There were a few paths of sex and order but they were of low magnitude (five effects for order, and two for sex out of 40 computed relationships) and they can be ignored for all practical purposes.

Intercorrelations and correlations with initiative

Table 6 displays the intercorrelations of the EOQ scales and their correlations with initiative (only Dutch language because initiative was only used in Dutch). Not unexpectedly, there were some high intercorrelations between the scales of EOQ. However, none of the correlations was so high that one would posit a duplication of content areas (even the $r = 0.57$ between error anticipation and error risk taking makes sense but does not signify overlapping construct content).

Correlations with initiative are in keeping with our hypotheses. Significant correlations appeared with error competence, learning from errors, error risk taking, negative with error strain, and negative with covering up. In addition, there were two correlations not hypothesized with thinking about errors and error communication.

Table 6. Intercorrelations of the EOQ scales and their correlations with initiative, Study II

	Error competence	Learning from errors	Error risk taking	Error strain	Error anticipation	Covering up errors	Communication about errors	Thinking about errors
Error competence								
Learning from errors	0.34*							
Error risk taking	0.21*	0.50*						
Error strain	-0.19†	0.01	-0.06					
Error anticipation	0.06	0.54*	0.57*	0.10				
Covering up errors	-0.10	-0.13	-0.17†	0.50*	-0.02			
Communication about errors	0.28*	0.51*	0.27*	-0.15	0.25*	-0.27*		
Thinking about errors	0.34*	0.51*	0.31*	0.17†	0.38*	-0.15	-0.27*	
Initiative	0.56*	0.35*	0.26*	-0.22*	0.08	-0.31*	0.29*	0.34*

* $p < 0.01$; † $p < 0.05$.

Overall discussion

Thus, we have succeeded in developing differentiated scales in the area of error orientation that are distinguishable from each other. (NB. we could prove this also with chi-square difference tests for each pair of constructs by testing a combined factor model against a separate factor model). While there are correlations, they are not too high. However, it may be that there are one or more higher-order constructs of error orientation, which would allow a short version of the questionnaire to be applied.⁶

The equivalence of the two versions in English and Dutch cannot be assumed for means but the two language versions correlate highly. Thus, there is some equivalence across languages. This means that mean differences should not be interpreted, but correlational differences can be interpreted.

We think the mean differences may have been caused by the translation. In general the items were less literally translated in the English than in the Dutch translation. While the English translators tried to find meaningful adaptations of the original German items, the wording became stronger because they changed conditionals or left out some of the moderating adverbs. As a result the subjects may have scored less on the English scales.

The lower internal consistencies in the English version of the EOQ still leave some open questions. But our design (comparison of two languages on one sample) allows us to exclude cultural differences or different experience as a reason for this finding. Instead, the fact that our subjects were not native English speakers seems to be a possible explanation.

A first test of validity has been made, producing the following plausible interpretations of the scales:

- (1) Error competence is active knowledge for immediate recovery from errors and reduction in error consequences. It relates to self-efficacy, to action-orientation after failure, need for achievement, and quite highly to initiative.

⁶ We could not test this hypothesis because in Study II the sample size was too small for such an analysis.

- (2) Learning from errors is the ability to prevent errors in the long term by learning from them, planning, and changing work processes. There are correlations with self-efficacy, qualification, plan-orientation, need for achievement, readiness to change, and initiative.
- (3) Error risk taking is the result of an achievement-oriented attitude which requires flexibility and taking responsibility. There are positive relations to need for achievement, qualification, readiness for change and initiative, as well as a negative relation to control rejection.
- (4) Error strain is characterized by a generalized fear of committing errors and by negative emotional reactions. It correlated negatively with self-efficacy, self-esteem, and initiative and positively with control rejection, psychosomatic complaints, depression, and negative affectivity.
- (5) Error anticipation is pessimistic and negatively tuned but at the same time it may be a realistic orientation. It correlated positively with negative affectivity and error strain, and negatively with optimism. The results of Study I are preliminary, since modeling of error anticipation was not perfect. Results of Study II show a clear correlation with learning from errors and with thinking about errors and with error risk taking. Possibly these results are similar to the well-known 'sadder but wiser' phenomenon (Alloy and Abramson, 1979).
- (6) Covering up errors is mainly the strategy of a non-self-assured person and may also be an adaptation to error-sensitive conditions at work, for example, job uncertainty. It relates to low self-esteem, negative affectivity, and high control rejection, and little initiative, but also to career stress and job uncertainty.

We think that these scales can have practical value. It may be useful to present a few speculations on how the EOQ can be related to safety strategies, to innovation (which necessarily implies that a number of errors have to be made), to change processes, to studies on errors, and to organizational culture.

New safety concepts might focus on error management (that is, basically reducing the negative error *consequences*, Frese, 1991) rather than on mere error prevention. Error management requires, for example, that personnel have well-developed anticipation and communication of errors, as well as error competence, to allow quick error detection and efficient error handling. The introduction of new safety strategies could be supported by the application of the EOQ. Reason (1990) has strongly advocated that organizational errors (or latent errors) are important pre-conditions for accidents. We think that some of what Reason means by latent errors (e.g. covering up errors, not taking them seriously, not anticipating them, etc.) may be related to issues measurable with the EOQ.

Innovation is closely related to initiative; the latter has been shown to be associated with an active and positive error orientation. We assume that fear of errors may often be a cause of resistance against change. Thus, the EOQ should be a useful instrument to analyze the potential for innovation processes.

When changes are introduced to increase responsibility of individuals (or groups) for their work process or product quality, one would assume that people with different error orientations will have different types of problems. Since higher responsibility and complexity at work implies that more errors appear, they may be covered up, they may not be detected if error anticipation is low, complexity may be avoided because people do not risk making errors, etc.

Experiments on errors could make use of the EOQ to combine interindividual differences with the effect of the experimental manipulation. For example error training, which proved to be a fruitful device to enhance learning (Frese, 1995) or studies on exploratory training (Dutke, 1994) should also deal with interindividual differences. Similarly, studies that observe errors that

appear in doing complex problem solving in simulated environments may usefully look at individual differences with the EOQ.

The EOQ may also be useful to measure the error culture of a group or a company. For measuring error culture, it may be useful to change the wording of the scales slightly and use 'we' or 'in our company' instead of 'I'.⁷ We think that an organization's error culture is determined by its members' orientation towards errors. In a 'positive' error culture there would be no punishment of errors, but a high orientation to learn from them and to actively deal with them, there would be high anticipation, communication and thinking about errors. As a consequence, there would be low error strain and covering up. In contrast, error prevention cultures with a high degree of bureaucracy would present the opposite picture. There may be differences in error orientation between different hierarchical levels of an organization, which could be an indicator of a problematic (overall) error culture (e.g. when higher echelons have a more constraining error culture than the lower echelons). Such discrepancies may also be the result of different degrees of complexity and control within one's job.

We assume that certain company cultural dimensions are related to different error cultures. For example, using Deal and Kennedy's (1982) categories, high risk companies will have higher risk taking but will only succeed if their error competences and learning from errors are high as well. Feedback speed will be related to error anticipation and we assume that companies with low feedback speed would do better if their error anticipation is high.

We also think that there are relationships to uncertainty avoidance (Hofstede, 1991). Cultures with high uncertainty avoidance should be low in risk taking, high in error strain, high in covering up errors, and low in error competence and learning from errors.

Obviously, the EOQ can and should be improved but we hope that it can have practical and theoretical value using these scales. We think that with the EOQ we have been able to produce a questionnaire that allows us to measure important issues related to errors.

Acknowledgements

The article was written with the support of the European Human Capital Mobility Grant (Prevention of Human Error in Systems for Energy Production and Process Industry, ERBCHRXCT930298). Study I was supported by the Deutsche Forschungsgemeinschaft (DFG, No Fr 638/6-5) (members M. Frese, D. Fay, S. Hilligloh, C. Speier, T. Wagner, J. Zempel). Study II was conducted in collaboration with the students: M. Admiraal, R. J. N. Clements, J. L. Hoekstra, D. J. Jacobson, S. P. Keijzer. We thank Frans Jeroen Oort for critiquing a previous version of this manuscript, Colin Oakley for improving the style of this paper and the unnamed reviewers for their helpful comments.

References

- Argyris, C. (1990). *Overcoming Organizational Defences: Facilitating Organizational Learning*, Prentice Hall, Englewood Cliffs, NJ.

⁷Van Dyck in our laboratory is currently using this approach to attempt to reveal error culture by applying the EOQ on the individual and organizational level. The appropriate wording can be obtained from the authors.

- Alloy, L. B. and Abramson, L. Y. (1979). 'Judgment of contingency in depressed and nondepressed students: Sadder but wiser?' *Journal of Experimental Psychology: General*, **108**, 441–485.
- Bainbridge, L. (1983). 'Ironies of automatization', *Automatica*, **19**, 775–779.
- Bandura, A. (1986). *Social Foundation of Thought and Action: A Social Cognitive Theory*, Prentice Hall, Englewood Cliffs, NJ.
- Bennett, M. (1977). 'Response characteristics of bilingual managers to organizational questionnaires', *Personnel Psychology*, **30**, 29–36.
- Bollen, K. A. (1989). *Structural Equations with Latent Variables*. Wiley, New York.
- Broadbent, D. E., Cooper, P. F., Fitzgerald, P. and Parkes, K. R. (1982). 'The Cognitive Failure Questionnaire (CFQ) and its correlates', *British Journal of Clinical Psychology*, **21**, 1–16.
- Brodbeck, F. C., Zapf, D., Prümper, J. and Frese, M. (1993). 'Error handling in office work with computers: A field study', *Journal of Occupational and Organizational Psychology*, **66**, 303–317.
- Browne, M. W. and Cudeck, R. (1993). 'Alternative ways of assessing model fit'. In: Bollen, K. A. and Long, J. S. (Eds) *Testing Structural Equation Models*, Sage Publications, London, pp. 136–162.
- Cronbach, L. J. and Meehl, P. E. (1955). 'Construct validity in psychological tests', *Psychological Bulletin*, **52**, 281–302.
- Cudeck, R. and Browne, M. W. (1983). 'Cross-validation of covariance structures', *Multivariate Behavioral Research*, **18**, 147–167.
- Deal, T. and Kennedy, A. (1982). *Corporate Cultures*, Penguin, Harmondsworth.
- Dormann, T. and Frese, M. (1994). 'Error training: Replication and the function of exploratory behavior', *International Journal of Human-Computer Interaction*, **6**, 365–372.
- Dutke, S. (1994). 'Error handling: Visualizations in the human-computer interface and exploratory learning', *Applied Psychology: An International Review*, **43**, 521–541.
- Fahrenberg, J. (1975). 'Die Freiburger Beschwerdenliste FBL', *Zeitschrift für Klinische Psychologie*, **4**, 79–100.
- Frese, M. (1984). *Do Workers Want Control or Don't They: Some Results on Denial and Adjustment*, IfHA-Berichte, Institut für Humanwissenschaft in der Arbeit und Ausbildung der Technischen Universität, Berlin.
- Frese, M. (1991). 'Error management or error prevention: Two strategies to deal with errors in software design'. In: Bullinger, H.-J. (Ed.) *Human Aspects in Computing: Design and Use of Interactive Systems and Work with Terminals*, Elsevier Science Publ., pp. 776–782.
- Frese, M. (1995). 'Error management in training: Conceptual and empirical results'. In: Bagnara, S., Zucchermaglio, C. and Stucky, S. (Eds) *Organizational Learning and Technological Change*, Springer, New York, pp. 112–124.
- Frese, M. and Plüddemann, K. (1993). 'Umstellungsbereitschaft im Osten und Westen Deutschlands', *Zeitschrift für Sozialpsychologie*, **24**(3), 198–210.
- Frese, M. and Zapf, D. (1991). 'Fehlersystematik und Fehlerentstehung: Eine theoretische Einführung'. In: Frese, M. and Zapf, D. (Eds) *Fehler bei der Arbeit mit dem Computer: Ergebnisse von Beobachtungen und Befragungen im Bürobereich*, Huber, Bern, pp. 14–31.
- Frese, M. and Zapf, D. (1994). 'Action as the core of work psychology: A German approach'. In: Triandis, H. C., Dunette, M. D. and Hough, L. M. (Eds) *Handbook of Industrial and Organizational Psychology*, Vol. 4, Consulting Psychologists Press, Palo Alto, CA.
- Frese, M., Erbe-Heinbokel, M., Grefe, J., Rybowski, V. and Weike, A. (1994). "'Mir ist lieber, wenn ich genau gesagt bekomme, was ich tun muß": Probleme der Akzeptanz von Verantwortung und Handlungsspielraum in Ost und West', *Zeitschrift für Arbeits- und Organisationspsychologie*, **38**, 22–33.
- Frese, M., Fay, D., Hilburger, T., Leng, K. and Tag, A. (1997). 'The concept of personal initiative: Operationalization, reliability, and validity in two German samples', *Journal of Occupational and Organizational Psychology*, **70**, 139–161.
- Frese, M., Kring, W., Soose, A. and Zempel, J. (1996). 'Personal initiative at work: differences between East and West Germany', *Academy of Management Journal*, **39**, 37–63.
- Frese, M., Stewart, J. and Hannover, B. (1987). Goal-orientation and planfulness: 'Action styles as personality concepts', *Journal of Personality and Social Psychology*, **52**, 1182–1194.
- Gerbing, D. W. and Hamilton, J. G. (1996). 'Viability of exploratory factor analysis as a precursor to confirmatory analysis', *Structural Equation Modeling*, **3**(1), 62–72.
- Grefe, J. (1994). *Managementfehler und Fehlermanagement: Eine explorative Studie*. Unveröffentlichte Diplomarbeit, Giessen.

- Hamilton, V. L. (1980). 'Intuitive psychologist process', *Journal of Personality and Social Psychology*, **39**, 767–772.
- Hofstede, G. (1991). *Cultures and Organizations*, McGraw-Hill, London.
- Jöreskog, K. G. and Sörbom, D. (1993). *LISREL 8: Structural Equation Modeling With the SIMPLIS Command Language*, Scientific Software International, Inc., Chicago.
- Kuhl, J. (1983). *Motivation, Konflikt und Handlungskontrolle*, Springer, Berlin.
- Lazarus, R. S. and Folkman, S. (1984). *Stress, Appraisal, and Coping*, Springer Publishing Co., New York.
- Leitner, K., Lüders, E., Greiner, B., Weber, W. G. and Hennes, K. (1987). *Analyse psychischer Belastung in der Arbeit. Das RHIA-Verfahren*, TÜV Rheinland, Cologne, Germany.
- McClelland, D. C. (1987). *Human Motivation*, Cambridge University Press, Cambridge.
- Modick, H. (1978). 'A three-scale achievement motive questionnaire: Report on a German modification of the Prestatie Motivatie Test', *German Journal of Psychology*, **2**, 8.
- Mohr, G. (1986). *Die Erfassung psychischer Befindensbeeinträchtigung bei Industriearbeitern*, Peter Lang, Frankfurt am Main.
- Oort, F. J. (1996). 'Using Restricted Factor Analysis in Test Construction', Doctoral thesis. University of Amsterdam.
- Peters, T. (1987). *Thriving on Chaos*, Harper & Row, New York.
- Poortinga, Y. H. (1989). 'Equivalence of cross-cultural data: An overview of basic issues', *International Journal of Psychology*, **24**, 737–756.
- Prümper, J., Zapf, D., Brodbeck, F. C. and Frese, M. (1992). 'Errors of novices and experts: Some surprising differences between novice and expert errors in computerized office work', *Behavior and Information Technology*, **11**, 319–328.
- Reason, J. T. (1990). *Human Error*, Cambridge University Press, New York.
- Scheier, M. F. and Carver, C. S. (1985). 'Optimism, coping and health: assessment and implications of generalized outcome expectancies', *Health Psychology*, **4**, 219–247.
- Schwarzer, R., Bäßler, J., Kwiatek, P., Schröder, K. and Zhang, J. X. (1997). 'The assessment of optimistic self-beliefs: Comparison of the German, Spanish, and Chinese versions of the general self-efficacy scale', *Applied Psychology: An International Review*, **46**(1), 69–88.
- Senge, P. M. (1990). *The Fifth Discipline: The Art and Practice of the Learning Organization*, Doubleday, New York.
- Semmer, N. (1984). *Streßbezogene Tätigkeitsanalyse*, Beltz, Weinheim, Germany.
- Speier, C. and Frese, M. (1996). 'Self-efficacy as a mediator between resources at work and personal initiative: A longitudinal field study in East Germany', manuscript, submitted for publication, University of Giessen.
- Stevens, J. (1992). *Applied Multivariate Statistics for the Social Sciences*, Lawrence Erlbaum, Hillsdale, NJ.
- Tanaka, J. S. (1993). 'Multifaceted conceptions of fit in structural equation models'. In: Bollen, K. A. and Scott, J. S. (Eds) *Testing Structural Equation Models*, Sage Publications, Newbury Park, CA, pp. 10–39.
- Watson, D., Clark, L. A. and Tellegan, A. (1988). 'Development and validation of brief measures of positive and negative affect: The PANAS-scales', *Journal of Personality and Social Psychology*, **54**, 1063–1070.
- Womack, J. P., Jones, D. T., Roos, D. and Sarpenter, D. S. (1990). *The Machine That Changed the World*, Rawson, New York.

Appendix

To what extent does this apply to you:

- 1 Not at all
 2 A bit
 3 Neither a bit, nor a lot
 4 A lot
 5 Totally

Items	R^2	
	Dutch	English
Error competence		
cpt1	When I have made a mistake, I know immediately how to correct it	0.44 0.27
cpt2	When I do something wrong at work, I correct it immediately	0.29 0.28
cpt3	If it is at all possible to correct a mistake, then I usually know how to go about it*	0.69 0.42
cpt4	I don't let go of the goal, although I may make mistakes	0.16 0.08
Learning from errors		
lrn1	Mistakes assist me to improve my work*	0.44 0.61
lrn2	Mistakes provide useful information for me to carry out my work*	0.45 0.61
lrn3	My mistakes help me to improve my work*	0.45 0.81
lrn4	My mistakes have helped me to improve my work	0.54 0.63
Error risk taking		
rsk1	If one wants to achieve at work, one has to risk making mistakes	0.45 0.49
rsk2	It is better to take the risk of making mistakes than to 'sit on one's behind'*	0.55 0.62
rsk3	To get on with my work, I gladly put up with things that can go wrong*	0.60 0.25
rsk4	I'd prefer to err, than to do nothing at all*	0.60 0.38
Error strain		
str1	I find it stressful when I err*	0.59 0.65
str2	I am often afraid of making mistakes	0.60 0.54
str3	I feel embarrassed when I make an error*	0.62 0.64
str4	If I make a mistake at work, I 'lose my cool' and become angry	0.56 0.28
str5	While working I am concerned that I could do something wrong	0.41 0.14
Error anticipation		
acp1	In carrying out my task, the likelihood of errors is high*	0.44 0.29
acp2	Whenever I start some piece of work, I am aware that mistakes occur	0.55 0.51
acp3	Most of the time I am not astonished about my mistakes because I expected them	0.35 0.34
acp4	I anticipate mistakes happening in my work	0.70 0.36
acp5	I expect that something will go wrong from time to time	0.46 0.28
Covering up errors		
cov1	Why mention a mistake when it isn't obvious?	0.50 0.40
cov2	It is disadvantageous to make one's mistakes public*	0.41 0.32
cov3	I do not find it useful to discuss my mistakes*	0.37 0.37
cov4	It can be useful to cover up mistakes	0.49 0.25
cov5	I would rather keep my mistakes to myself*	0.55 0.64
cov6	Employees who admit to their errors, make a big mistake	0.37 0.31
Error communication		
com1	When I make a mistake at work, I tell others about it in order that they do not make the same mistake	0.24 0.22
com2	If I cannot rectify an error by myself, I turn to my colleagues*	0.62 0.53
com3	If I cannot manage to correct a mistake, I can rely on others*	0.26 0.24
com4	When I have done something wrong, I ask others, how I should do it better	0.45 0.50
Thinking about errors		
thk1	After I have made a mistake, I think about how it came about	0.45 0.43
thk2	I often think: 'How could I have prevented this?'	0.38 0.36
thk3	If something goes wrong at work, I think it over carefully	0.59 0.66
thk4	After a mistake has happened, I think long and hard about how to correct it*	0.44 0.33
thk5	When a mistake occurs, I analyze it thoroughly*	0.72 0.76

* New item in Study II.