ERGONOMICS, 1995, VOL. 38, NO. 7, 1456–1467

Cognitive failures and performance differences: validation studies of a German version of the cognitive failures questionnaire

P. L. KLUMB

Department of Psychology, Berlin University of Technology, Dovestr. 1-5, D-10587 Berlin, Germany

Keywords: Error propensity; Questionnaire; Stress vulnerability.

A German version of Broadbent's Cognitive Failures Questionnaire was developed because there was no comparable German instrument. As external validity criteria observable failures in everyday situations were used. The positive correlations were significant but moderate in size. For construct validation, inventories of trait anxiety, coping styles and action control were used. Most of the correlations were significant and in the expected direction. A maximum likelihood factor analysis of the questionnaire suggested that the utility of the total score may be doubtful. The unidimensionality of the construct requires further investigation.

1. Introduction

A large proportion of accidents and incidents within a broad range of industries is attributed to 'human error'. Depending on the definition of 'error' and the methods of data collection and analysis that are used, estimates range from 44% to 92% (Fegetter 1982, INPO 1984, Perrow 1984, Swain and Guttmann 1983). Analyses of the causes of errors and accidents have resulted in a controversial discussion concerning the question of whether there are individuals who are more error prone than others and might, therefore, be more often involved in accidents.

The concept of 'accident proneness' was already used at the beginning of this century (Selz 1919). However, both this early and later differential approaches to accident prevention have been heavily criticized on conceptual and methodological grounds (cf. Hacker 1986, Hoyos and Zimolong 1988). Applied to errors instead of accidents the differential approach is considered to be more adequate because errors are related in a direct way to the respective activity (Leplat 1985, p. 142). However, in spite of the implicit equation of errors with accidents in research on safety their homogeneity cannot be assumed (Hale and Glendon 1987, Wehner and Stadler in press).

In general, errors cannot be exclusively attributed to personality characteristics but are a consequence of a mismatch between the task, the technical system and the human or social system (for more on the concept of mismatch, cf. Rasmussen 1987, Zapf *et al.* 1989). This incompatibility manifests itself in a deviation from a specified goal. Therefore, strategies of error prevention and management can be aimed at the technical system (i.e., (re)design measures), at the administrative system (i.e., organizational design) or at the social system (including personnel training and selection).

To shed light on the role of individual predispositions conducive to errors, studies that examine the construct of error proneness and its relationships to other constructs as well as performance data are required. Recent research by Broadbent and co-workers

and by Reason and co-workers has gone in this direction and their approach will be outlined in the following section.

2. Cognitive control

Broadbent *et al.* (1982) assume that diverse perceptual, memory and action failures are influenced by a global and rather stable factor that is relatively independent of traditional personality and intelligence measures. The general term 'cognitive failures' was chosen to cover all types of execution failures and/or storage failures because of their common origin in the cognitive organization of the individual, while excluding failures of ability or planning. According to Reason (1988, 1990) the variety of failures observed at the skill-based level of performance can be mainly attributed to two kinds of control deficits in automatic processing: inattention and overattention. The forms that these control-mode failures can take are double capture slips, omissions following interruptions, reduced intentionality, perceptual confusions and interference errors in the case of inattention; omissions, repetitions and reversals in the case of overattention (see Reason (1990) for the GEMS framework, i.e., the Generic Error Modelling System). These defective control processes seem to be tied to the way in which people deploy their limited attentional resources. An inflexible style of attentional resource management may lead to a propensity to everyday slips and lapses, as well as to enhanced stress vulnerability (because of inefficient coping).

There are two main operationalizations of this cognitive style. Both are based on the self-reported frequency of everyday errors as measured by standardized questionnaires about memory and metamemory: the Error Proneness Questionnaire by Reason and Mycielska (1982) and the Cognitive Failures Questionnaire (CFQ) by Broadbent *et al.* (1982). As these measures seem to be highly correlated, the CFQ was chosen for the present research because there is more evidence existing from field and laboratory studies.

The original version of the CFQ consisted of 25 items that represented failure modes at the skill-based level (presumably) occurring to everybody from time-to-time and were intended to cover the areas of perception, memory and action. For each item, respondents were asked to indicate the frequency of failures happening to him or her during the last six months. All items had to be answered on five-point scales ranging from never to very often. The questions were worded in a positive direction and it took about ten minutes to answer them. Total CFQ scores were calculated by summing across item ratings (0 to 4) for each subject.

For the purpose of validation of the questionnaire, Broadbent and his colleagues used the judgements of reference persons (Broadbent *et al.* 1982) and performance parameters assessed in tasks that represented different attentional demands (Broadbent *et al.* 1986). However, a clear relationship between task performance and CFQ scores still remains to be shown. Although individual differences in multiple-task performance (in the laboratory) were discussed within the domain of workload research as reflecting different strategies or cognitive styles (Damos *et al.* 1983) real-life situations seem to be more appropriate for revealing the relevant differences in cognitive control (cf. Hermann 1982: lack of scale to task isomorphism).

While Larson and Merritt (1991) found a relationship between CFQ score and the frequency of involvement in traffic accidents for young US army recruits, studies on the effects and after-effects of working at visual display units (VDU; Smith and Peck 1989, Smith *et al.* 1989) found no effect of CFQ scores on performance. However, for people with high CFQ scores work at the VDU had a negative effect on subjective

Table 1. Description of the samples of the studies.

	Number of subjects	Number of female subjects	Age (years)	
			Mean	Range
Sample 1				
Library	102	43	27	18–57
Dry cleaners	37	19	33	17–58
Lost property office	37	9	39	16–82
Sample 2				
Students	74	0	22	19–28
Sample 3				
(Library, dry cleaners and students) Item and test statistics	213	62	25	17–58

well-being. Furthermore, these subjects seemed to be more prone to develop negative long-term effects (i.e., neurotic and somatic symptoms).

Since the relationship between the propensity for cognitive failures assessed by questionnaires and performance differences is thus still controversial, the aim of the present study was a further examination of the construct. Three issues were addressed. First, the relationships of the CFQ to clearly specified failures in everyday performance were investigated; second, the association of the CFQ with theoretically related psychological constructs was looked at; and third, the proposed unidimensionality of the cognitive failures questionnaire was examined.

3. Samples

The characteristics of the samples of the three studies are summarized in table 1. It contains the number of participants, their age and gender. In the criterion-related validation study the subjects of the three settings are described separately; within the different settings, the groups were homogeneous with regard to age. For the construct-related validation the sample consisted of 74 male students of the faculty of electrical engineering who completed the questionnaires in a lecture room of the Berlin University of Technology. For the computation of item and scale statistics the total 213 questionnaires were pooled. The subjects of the lost property office sample were excluded from the analyses because several inconsistencies in their responses suggested that they might have been rather defensive.

4. Study 1: Criterion-related validation

4.1. *Method*

For the first German version of the CFQ the items, wording, response format and instruction from the original questionnaire were adopted. As precise agreement in the wording of the items was not of paramount importance, it was thought to be sufficient to guarantee the understandability of the German items by having the translation checked by several colleagues and a student sample. The result of this process was a revised version of the questionnaire, in which the question format of the items was replaced by a statement format and the lower end of the response scale was changed

from 'never' to 'hardly ever' because people were found to be reluctant to use the 'never' option. Seven further items were added to enhance the representation of different areas of cognitive failure (these CFQ items are listed in the Appendix).

The aim of the first study was to combine the testing of the German version of the CFQ with a first criterion-related validation. The hypothesis to be tested was that people who have higher scores in the cognitive failures inventory—presumably indicative of a defective management of cognitive resources—are more liable to certain everyday problems such as forgetting to return borrowed books in time or losing things in public places. Three settings were selected in which the consequences of the above-mentioned problems could be observed with particular frequency, and their authors could be questioned: libraries, dry cleaners, and a lost property office. A quasi-experimental design was used within each of these settings. The CFQ score was determined for those clients who returned books late, who tried to get back their clothes at the dry cleaners without a ticket or who tried to reclaim property; these persons constituted the experimental groups. Ideally, the assessment of the self reports should be remote in time and space from the commitment of the failure. However, research-economical considerations lead to the decision to assess them concurrently. This procedure may have sensitized the subjects of the experimental groups and therefore may not meet the non-reactivity postulate. At the same times and locations individuals who did not show the behaviour in question were assigned randomly to the control groups (in the lost property office, these were people who reported to be present on behalf of somebody else). As a manipulation check, three further items were added to the questionnaire that asked for the frequency of the target failures (e.g., returning books to the library after the deadline passed).

4.2. *Results*

The answers to the questions concerning the frequency of returning books late, forgetting the ticket from the dry cleaners or losing an object were cross-tabulated to examine, separately for each setting, whether the respective target failure happened more frequently to the subjects of the experimental group compared to those of the control group. The reactivity objected to above may have affected the validity of the manipulation-check procedure. The results of the manipulation checks are shown in table 2 for the experimental and control groups for the three settings.

Except for the lost property office group, the experimental group indeed reported more failures of the type that had caused them to be approached. Chi-square was used for comparing proportions in the 2×2 contingency tables consisting of the frequencies of the dichotomized control items. In order to avoid zero cells categories 1 and 2 (hardly ever and quite rarely) had been pooled and compared to the pooled categories 3 to 5. Including the Yates correction the tests yielded the following results: within the lost property office sample there was no difference between experimental and control group. Within the library group a Chi-square of 16·71 (d.f. = 1, $\alpha < 0{\cdot}001$) and for the distribution within the dry cleaning sample a Chi-square of 15·67 (d.f. = 1, $\alpha < 0{\cdot}001$) were obtained. Hence, the assignment of subjects to the experimental vs. control groups by means of our performance criterion was warranted.

With regard to the CFQ (both the total score and, with two exceptions, the individual items), the experimental group scored higher than the control group within each separate setting, as well as in an overall comparison. The relationship between total score and dichotomous criterion (experimental group coded as 1, control group coded as 0) in the total group was expressed by a point biserial correlation coefficient of 0·20 for the

Table 2. Distribution of answers to the control questions for experimental and control groups in the respective settings.

	Hardly ever	Quite rarely	Occasionally	Quite often	Very often
Library groups					
Experimental	1	9	9	12	13
	2·3%	20·5%	20·5%	27·3%	29·5%
Control	19	19	13	6	1
	32·8%	32·8%	22·4%	10·3%	1·7%
Dry cleaning groups					
Experimental	0	4	4	4	2
		28·6%	28·6%	28·6%	14·3%
Control	15	7	0	0	1
	65·2%	30·4%			4·3%
Lost property office groups					
Experimental	1	15	0	2	0
	5·6%	83·3%		11·1%	
Control	8	8	1	1	0
	44·4%	44·4%	5·6%	5·6%	

original version of the questionnaire and of 0·22 for the enlarged version developed for this study. Both coefficients are significant at the 1% level.

4.3. *Discussion*

The correlations obtained between the CFQ score and specified failures in everyday situations were in the expected directions and significant, but only of modest size. However, bearing in mind that the criterion merely consists of one target behaviour representing the equivalent of one single item a coefficient of 0·20 can be judged to be quite substantial.

5. Study 2: Construct validation

5.1. *Method*

In addition to the correlation with external criteria, an attempt to construct-validate the questionnaire was made. Based on theoretical considerations, the following constructs were selected: anxiety, assessed with the German trait version of the anxiety inventory STAI (Laux *et al.* 1981); ways of coping as operationalized in the 42-item version of the coping inventory SVF (Janke *et al.* 1990) and action control after success, failure and in prospect measured by the questionnaire HAKEMP90 (Kuhl 1990) (see table 3 for a summary of the scales).

The STAI was included to represent one of the traditional and commonly administered personality measures. A good correlation with the CFQ was not expected (Broadbent *et al.* found a modest correlation in 1982, while reporting a moderating effect of anxiety in 1986).

Action control *sensu* Kuhl is a construct that is closely related to cognitive control/cognitive failure. The three subscales of the HAKEMP90 consist of 12 items each. The HOM scale represents action orientation after failure, HOP represents action orientation in the course of planning an action, and HOT refers to action orientation during the execution of an activity. The scales were expected to be negatively correlated

Table 3. Means, standard deviations, number of items, range and reliabilities of the scales used for construct validation.

Scale	Mean	Standard deviation	Number of items	Range	Cronbach's α
CFQ	1·49	0·41	32	0–4	0·82
CFQ (original)	1·38	0·37	25	0–4	0·78
STAI	40·07	8·80	20	1–4	0·89
SVF					
Inefficient coping	11·51	3·90	8	0–4	0·62
HAKEMP					
HOM	5·50	2·66	12	0–1	0·67
HOP	4·62	2·75	12	0–1	0·72
HOT	8·30	2·05	12	0–1	0·46

with the CFQ score because the inefficient style of action control is represented by smaller scores. The CFQ item about making up one's mind was expected to correspond well with the HOP scale.

The SVF contains 21 subscales consisting of two items each, which represent different coping strategies such as seeking social support. The probability of adopting each strategy in case of 'being disturbed, getting agitated or losing one's balance' had to be indicated. There was no differentiation with regard to the type of stressor encountered.

The propensity for cognitive failures is thought to be linked to stress vulnerability through the higher order mechanism of attentional resource management. Therefore, it was hypothesized that high CFQ scores would be associated with adopting maladaptive coping strategies. Since the efficacy or efficiency of coping strategies can only be judged in relation to situational characteristics (cf. Parkes cited in Broadbent *et al.* 1986) a scale was sought that would represent those ways of coping that are ineffective in most circumstances.

Kühlmann (1990), working with an adapted SVF version, found a factor called 'unstable submission' in his study of the stress and coping behaviour of urban bus drivers. This factor was described as resigned acceptance together with a propensity for emotional outbursts, and it was the only one of four factors that showed a relationship with negative stress effects. The unstable submission factor was represented by the subscales 'resignation' and 'aggression' of the SVF, which were collapsed in the present study with the subscales 'continued mental occupation with the stressor' and 'drug use' to the scale 'inefficient coping strategies'. For this scale a positive relationship to CFQ was expected.

Following a recommendation of Kühlmann (1990), an index of the range of available coping strategies was computed additionally. This range was operationalized by the number of coping strategies with a probability of being used greater than or equal to 'possibly' (score of 2 on the 5-point scale ranging from never (0) to very likely (4)). A negative correlation between CFQ and the range of available strategies was expected. High CFQ scorers—characterized by an inflexible way of deploying their cognitive resources—were assumed to have a limited repertoire of coping strategies.

An additional item was included at the end of the CFQ asking the subjects to indicate on a five-point scale whether they were particularly stressed during the last months. This item was expected to be uncorrelated to the CFQ score.

Table 4. Interrelations between the total score of the CFQ and six scales of the selected instruments.

	COP-range	COP-neg	STAI	HOM	HOT	HOP
CFQ-32	0·31*	0·49**	0·45**	−0·25	−0·16	−0·60**
CFQ-25	0·29*	0·49**	0·42**	−0·26	−0·17	−0·57**

$n = 74$.
* Significant, $p < 0·05$.
** Significant, $p < 0·01$.

5.2. *Results*

The relationships of the CFQ to the selected instruments were computed for both the 25- and the 32-item version of the CFQ and are given in table 4. Nearly all of the correlations between the computed scales and the CFQ total score were significant and in the expected directions, i.e., most of the hypothesized associations of the cognitive-failure construct with the selected constructs were confirmed: both total CFQ scores were found to have a substantial and significant positive correlation with the scale 'inefficient coping strategies' and the correlations to the HAKEMP scales were negative but reached significance only in the case of the HOP (action orientation in the course of planning an action). This finding is not surprising since the HOP scale represents the more general aspects of cognitive organization. The additionally computed correlation between the CFQ item 15 (trouble making up one's mind) and the HAKEMP scale HOP was −0·27; the coefficient being significant at the 5% level. In contrast to our expectations the relationship between CFQ score and the computed range of coping strategies turned out to be positive. A significant positive correlation between CFQ and STAI was found that was unexpected with regard to its size. As expected, the amount of stress experienced in the course of the past months was not correlated to the CFQ score ($r = -0·02$). However, a coefficient of 0·38 ($p < 0·01$) was found between the stress item and the STAI score. The representativeness of the sample was checked by comparing the intercorrelations between the STAI-trait scale and the SVF subscales with those reported in the STAI manual (Laux *et al.* 1981). As to those subscales that are comparable across the different SVF versions, the coefficients had a similar size.

5.3. *Discussion*

The correlations found between the CFQ score and the measures of action control and coping strategies show that high and low CFQ scorers differ in their action control, in their preferred ways of coping, as well as in their range of available coping strategies.

The pronounced correlation between CFQ and STAI was not in accord with our expectations. However, Matthews and Wells (1988) obtained a comparable correlation. They showed that this positive correlation between cognitive failures and anxiety was mediated by the personality trait of self-consciousness. Inefficient processing of information in anxious persons is generally attributed to negative self-evaluations and worries. The processing of cognitions about the self takes up attentional capacity and is hence an explanation for both the disruption in performance and the adoption of maladaptive coping strategies.

The picture is finished by the correlation between the stress item and the CFQ score. This again is support for the assumption that it is not increased stress that results in a higher frequency of cognitive failures. The whole pattern of correlations can be

interpreted as further evidence in accord with the assumption of a higher order mechanism controlling various specific functions (Broadbent *et al.* 1982, Norman and Shallice 1985, Reason 1984).

6. Study 3: Consistency, homogeneity and factor structure

6.1. *Method*

On the basis of 213 data sets pooled from the student sample and the subjects of the libraries and dry cleaners, analyses of the consistency, homogeneity and factor structure of both the 32-item version and the 25-item version of the CFQ were conducted.

With regard to factor analytic procedures the sample size met the criterion of at least 200 subjects (Thorndike 1978) but failed to meet the criterion of ten times as many subjects as items (Nunnally 1978). Following a recommendation of Tobias and Carlson (1969), Bartlett's test of sphericity was applied to test whether the variables were correlated in the population. The obtained Chi-square of 822·76 was significant ($p = 0{\cdot}00000$). Hence it was possible to factorize the correlation matrix. A principle component analysis (PCA) was applied as an exploratory procedure to identify the number of dimensions of the 32-item version of the CFQ.

Parallel analysis (Horn 1965, Humphreys and Ilgen 1969) was chosen to determine the number of components to retain. This technique compares the eigenvalues obtained from the empirical data with those generated from a set of normally distributed random numbers with the same number of items and observations. According to Zwick and Velicer (1986) this procedure is sufficiently accurate to be useful and not greatly affected by data parameters.

6.2. *Results*

All items tended to be positively correlated, which was comparable to the findings reported by Broadbent *et al.* 1982, 1986. The mean, standard deviation and corrected item-total correlation of each question are shown in table 5.

Cronbach's coefficient α was 0·82 (0·78 for the 25-item version) and the split-half coefficient computed according to the Spearman-Brown formula was 0·83 (Broadbent *et al.* 1982 reported an α of 0·89).

The initial PCA extracted 11 factors with eigenvalues greater than 1. According to the Horn test two components were retained because their eigenvalues lay above the highest one of the random sample with a value of 1·82. The two-factor solution was then realized with the aid of a maximum likelihood factor analysis with varimax rotation. The first factor (eigenvalue = 3·63) accounted for 13·3% of the variance and was labelled *distractibility*. The leading items loading on this factor were the following: CFQ30 and CFQ17 (forget the location of things), CFQ21 (get distracted), CFQ26 (forget to mention an aspect in a conversation), CFQ16 (forget appointments), CFQ22 (tip-of-the-tongue phenomenon), CFQ19 (daydream when you ought to be listening), CFQ28 (forget to transfer a message) and CFQ20 (forget people's names). The second factor (eigenvalue = 2·24) accounted for 5·1% of the variance and was named *sensori-motor coordination*. It gathered all those items referring to motor slips: CFQ31 (bump into objects), CFQ32 (stumble), CFQ24 (drop things) and CFQ5 (bump into people). A further item of this factor CFQ10 (lose one's temper), may be an indicator of an impulsivity component. In the light of the factor-analytic findings, the results of the construct validation studies were reanalysed. Not a single significant correlation was

Table 5. Item statistics ($n = 213$).

Item	Mean	Standard deviation	Corrected item-total correlations	Percentage of answers ⩾ 3
CFQ1	2·24	0·95	0·33	39
CFQ2	1·55	1·05	0·41	17
CFQ3	1·05	1·03	0·38	11
CFQ4	0·93	1·18	0·20	12
CFQ5	1·07	0·95	0·30	8
CFQ6	1·41	1·16	0·24	19
CFQ7	1·90	1·15	0·33	34
CFQ8	1·78	1·16	0·26	25
CFQ9	1·60	1·06	0·20	20
CFQ10	0·99	1·00	0·22	9
CFQ11	1·56	1·42	0·30	33
CFQ12	0·59	0·88	0·24	5
CFQ13	1·62	1·12	0·35	27
CFQ14	1·09	1·03	0·23	12
CFQ15	2·28	1·13	0·19	46
CFQ16	0·71	0·88	0·40	6
CFQ17	2·08	1·18	0·44	40
CFQ18	0·43	0·72	0·26	3
CFQ19	2·60	1·00	0·44	55
CFQ20	2·23	1·11	0·32	39
CFQ21	2·11	1·07	0·46	37
CFQ22	1·57	1·07	0·36	19
CFQ23	0·47	0·76	0·31	3
CFQ24	1·47	1·05	0·40	17
CFQ25	1·20	1·02	0·42	10
CFQ26	1·94	0·94	0·41	27
CFQ27	1·96	1·12	0·31	34
CFQ28	1·39	1·08	0·32	17
CFQ29	1·55	1·08	0·30	19
CFQ30	1·74	1·07	0·46	22
CFQ31	1·38	1·02	0·39	15
CFQ32	0·96	0·88	0·23	5

found between the factor *sensori-motor coordination* and the scales used for construct validation.

6.3. *Discussion*

The results of the factor analyses were not in accordance with Broadbent's finding of a general cognitive-failure factor (Broadbent *et al.* 1982). However, in line with this evidence were the findings of Porter and Corlett (1989) that the correlations between the total scores of the CFQ and their own accident proneness questionnaire (APQ), which focuses on motor slips (e.g., stumble, cut or bruise oneself, drop things), as well as between the CFQ and the scores on several performance tests, were insignificant. Only one CFQ item was found to discriminate between two groups formed on the basis of the APQ, and this item was identical to one of the APQ items. In addition Pollina *et al.* (1992), using the original CFQ version ($n = 387$), and Miller (1991), using a Hungarian version of the 32-item CFQ form ($n = 282$), extracted more than one component.

Factor analysis on small samples may not be very reliable and the structure of the CFQ remains to be confirmed (particularly the number of components to be retained, cf. Matthews *et al.* 1990). However, it must be seriously doubted whether the CFQ score is a useful predictor of specific performance deficits. It may be necessary to distinguish between different subscales, which are presently not all that reliable.

7. Discussion and conclusions

Minor errors assessed in three everyday situations (libraries, dry cleaners, and lost property office) were related to the scores of the Cognitive Failures Questionnaire. The obtained coefficients were significant and in the expected directions, but they were only moderate in size. Parallel findings were also obtained for the associations of the CFQ with relevant psychological constructs. This pattern of interrelations seems to be theoretically interesting and worthy of being further investigated.

One explanation for the low coefficients might be the component structure of the CFQ, which was characterized by a separation between items concerning attentional focus and sensori-motor coordination. Evidently, and in accordance with the findings of Porter and Corlett (1989), the correlations with external criteria could be enhanced by using appropriate subscales instead of the total score. Perhaps, therefore, the idea of a homogenous cognitive-failure dimension has to be abandoned in favour of more specific constructs. Since the dimensionality of the CFQ has theoretical as well as practical diagnostic implications, further studies are necessary to verify the identified correlation patterns.

The advantage of self-report questionnaires is that they are easy to administer and that the information obtained comes directly from the source. One of their main disadvantages is that it is 'possible for respondents to project an incorrect self-image of themselves' (Porter and Corlett 1989, p. 320). In the case of the CFQ, several factors may foster this tendency. First of all, defensiveness may play an important role as people may be reluctant to frankly report their weaknesses. People tend to present a favourable impression of themselves especially in situations with an inherent potential of assessment and contingent sanctions. Alternatively, people may simply have an inaccurate view of themselves. In particular the reliability of the metamemory regarding statements about memory functioning may be questioned. Failures may remain undetected or they may be forgotten after some time. Hence the usefulness of failure questionnaires for selection purposes is doubtful, especially when they possess a high degree of transparency of face validity. The assessment of performance indices would be preferable. For example, Smith and Peck (1989) found proof-reading speed to be a useful index, which in addition would be easy to assess. The studies of Porter and Corlett (1989) indicated that task complexity may be a crucial characteristic that should be examined more closely in further research.

Finally, a critical remark concerning the methods of item collection for error-proneness questionnaires should be made. The inventory developed by Porter and Corlett (1989) was 'the product of a brainstorming session with 15 colleagues' (p. 319). Broadbent *et al.* in constructing the CFQ also relied on information from 'ourselves and acquaintances' (1982, p. 3). This way of item generation may emphasize cultural specificities of a certain group of people that are not shared by other groups. For instance, although within the questionnaires it is stressed that they are about minor failures that happen to *everybody*, many of the subjects after answering the questions, asked whether the purpose was to turn them into 'fools'. Committing little errors apparently does not have the same appeal to everybody and a self-report measure may

reflect both the subjects' sensitivity and their level of aspiration with regard to their error production. For that reason, it is suggested that in future research, different samples should rate the frequency of each of the described events for a hypothetical average person to test the general plausibility of the items.

Acknowledgements

Thanks are due to Donald Broadbent, Dieter Zapf and two anonymous reviewers for their comments and last, but not least to the subjects for their participation. The studies were conducted in conjunction with 12 students in the course of a student project at the Berlin University of Technology.

References

BROADBENT, D. E., BROADBENT, M. H. P. and JONES, J. L. 1986, Performance correlates of self-reported cognitive failure and of obsessionality, *British Journal of Clinical Psychology*, **25**, 285–299.

BROADBENT, D. E., COOPER, P. F., FITZGERALD, P. and PARKES, K. R. 1982, The Cognitive Failures Questionnaire (CFQ) and its correlates, *British Journal of Clinical Psychology*, **21**, 1–16.

DAMOS, E. L., BITTNER, A. C. and SMIST, T. E. 1983, Individual differences in multiple-task performance as a function of response strategy, *Human Factors*, **25**, 215–226.

FEGETTER, A. J. 1982, A method for investigating human factor aspects of aircraft accidents and incidents, *Ergonomics*, **25**, 1065–1075.

HACKER, W. 1986, *Arbeitspsychologie* (Deutscher Verlag der Wissenschaften, Berlin).

HALE, A. R. and GLENDON, A. I. 1987, *Individual Behaviour in the Control of Danger* (Elsevier, Amsterdam).

HERRMANN, D. J. 1982, Know thy memory: the use of questionnaires to assess and study memory, *Psychological Bulletin*, **92**, 434–452.

HORN, J. L. 1965, A rationale and test for the number of factors in factor analysis, *Psychometrika*, **30**, 179–185.

HOYOS, C. G. and ZIMOLONG, B. 1988, *Occupational Safety and Accident Prevention: Behavioral Strategies and Methods* (Elsevier, Amsterdam).

HUMPHREYS, L. G. and ILGEN, D. R. 1969, Note on a criterion for the number of common factors, *Educational and Psychological Measurement*, **29**, 571–578.

INPO 1984, An analysis of root causes in 1983 Significant Event Reports, *INPO 84-027* (Institute of Nuclear Power Operations, Atlanta, GA).

JANKE, W., ERDMANN, G. and KALLUS, W. 1990, Streßverarbeitungsfragebogen (SVF) Kurzformen, Beschreibung und Handanweisung, Unpublished manual, Department of Psychology, University of Würzburg.

KUHL, J. 1990, Kurzanweisung zum Fragebogen HAKEM90 (Handlungskontrolle nach Erfolg, Mißerfolg und prospektiv), Department of Psychology, University of Osnabrück.

KÜHLMANN, T. M. 1990, Coping with occupational stress among urban bus and tram drivers, *Journal of Occupational Psychology*, **63**, 89–96.

LARSON, G. E. and MERRITT, C. R. 1991, Can accidents be predicted? An empirical test of the Cognitive Failures Questionnaire, *Applied Psychology: An International Review*, **40**, 37–45.

LAUX, L., GLANZMANN, P., SCHAFFNER, P. and SPIELBERGER, C. D. 1981, *Das State-Trait-Angstinventar STAI* (Beltz, Weinheim).

LEPLAT, J. 1985, *Erreur humaine, fiabilité humaine dans le travail* (Armand Colin, Paris).

MATTHEWS, G. and WELLS, A. 1988, Relationships between anxiety, self-consciousness, and cognitive failure, *Cognition and Emotion*, **2**, 123–132.

MATTHEWS, G., COYLE, K. and CRAIG, A. 1990, Multiple factors of cognitive failure and their relationships with stress vulnerability, *Journal of Psychopathology and Behavioral Assessment*, **12**, 49–65.

MILLER, R. 1991, Factor structure of a German and a Hungarian version of the Cognitive Failures Questionnaire, Unpublished manuscript, Berlin University of Technology.

NORMAN, D. A. and SHALLICE, T. 1985, Attention to action: Willed and automatic control of behaviour, in R. J. Davidson, G. E. Schwartz and D. Shapiro (eds), *Consciousness and Self-regulation: Advances in Research*, Vol. 14 (Plenum Press, New York).

NUNNALLY, J. C. 1978, *Psychometric Theory* (MacGraw-Hill, New York).
PERROW, C. 1984, *Normal Accidents* (Basic Books, New York).
POLLINA, L. K., GREENE, A. L., TUNICK, R. H. and PUCKETT, J. M. 1992, Dimensions of everyday memory in young adulthood, *British Journal of Psychology*, **83**, 305–321.
PORTER, C. S. and CORLETT, E. N. 1989, Performance differences of individuals classified by questionnaire as accident prone or non-accident prone. *Ergonomics*, **32**, 317–333.
RASMUSSEN, J. 1987, The definition of human error and a taxonomy for technical system design, in J. Rasmussen, K. Duncan and J. Leplat (eds), *New Technology and Human Error* (Wiley, Chichester).
REASON, J. 1984, Lapses of attention in everyday life, in R. Parasuraman and D. R. Davies (eds), *Varieties of Attention* (Academic Press, London), 515–549.
REASON, J. 1988, Stress and cognitive failure, in S. Fisher and J. Reason (eds), *Handbook of Life Stress, Cognition and Health* (Wiley, Chichester).
REASON, J. 1990, *Human Error* (Cambridge University Press, New York).
REASON, J. and MYCIELSKA, K. 1982, *Absent-Minded? The Psychology of Mental Lapses and Everyday Errors* (Prentice-Hall, Englewood Cliffs).
SELZ, O. 1919, Über den Anteil der individuellen Eigenschaften der Flugzeugführer und Beobachter an Fliegerunfällen, Eine psychologische Untersuchung auf unfallstatistischer Grundlage, *Zeitschrift für angewandte Psychologie*, **15**, 254–296.
SMITH, A. P. and PECK, D. 1989, The influence of time of day and cognitive failure on the effects of working at a visual display unit, in D. Brogan (ed.), *Proceedings of the First International Conference on Visual Search* (Taylor & Francis, London), 345–352.
SMITH, A. P., PECK, D. and CLATWORTHY, T. 1989, After-effects of working at visual display units, *Work & Stress*, **3**, 195–201.
SWAIN, A. D. and GUTTMANN, H. E. 1983, *Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications* (Nuclear Regulatory Commission, Washington, DC).
THORNDIKE, R. M. 1978, *Correlational Procedures for Research* (Gardner Press, New York).
TOBIAS, S. and CARLSON, J. E. 1969, Brief report: Bartlett's test of sphericity and chance findings in factor analysis, *Multivariate Behavioral Research*, **4**, 375–378.
WEHNER, T. and STADLER, M. (in press), Gestaltpsychologische Beiträge zur Struktur und Dynamik fehlerhafter Handlungsabläufe, in H. Heckhausen and J. Kuhl (eds), *Motivation, Volition und Handlung. Band CIVI4 der Enzyklopädie der Psychologie* (Hogrefe, Göttingen).
ZAPF, D., BRODBECK, F. and PRÜMPER, J. 1989, Handlungsorientierte Fehlertaxonomie in der Mensch-Computer Interaktion, *Zeitschrift für Arbeits- und Organisationspsychologie*, **33**, 178–189.
ZWICK, W. R. and VELICER, W. F. 1986, Comparison of five rules for determining the number of components to retain, *Psychological Bulletin*, **99**, 432–442.

Appendix. Seven items added to the CFQ

26. At the end of a conversation I realized that I forgot to mention an aspect which I had intended to bring in.
27. After leaving my apartment I had to return in order to pick up something that I failed to take with me.
28. I forgot to transfer a message to somebody as I was requested to do.
29. I could not remember something that I had been told some time ago.
30. I could not find something that I only put away a couple of days ago.
31. I bumped against an object.
32. I stumbled on the street.