

The Cognitive Failures Questionnaire (CFQ) and its correlates

D. E. Broadbent, P. F. Cooper, P. FitzGerald and K. R. Parkes

This paper describes a questionnaire measure of self-reported failures in perception, memory, and motor function. Responses to all questions tend to be positively correlated, and the whole questionnaire correlates with other recent measures of self-reported deficit in memory, absent-mindedness, or slips of action. The questionnaire is however only weakly correlated with indices of social desirability set or of neuroticism. It is significantly correlated with ratings of the respondent by his or her spouse, and accordingly does have some external significance rather than purely private opinion of the self. The score is reasonably stable over long periods, to about the same extent as traditional measures of trait rather than state. Furthermore, it has not thus far been found to change in persons exposed to life-stresses. However, it does frequently correlate with the number of current psychiatric symptoms reported by the same person on the MHQ; and in one study it has been found that CFQ predicts subsequent MHQ in persons who work at a stressful job in the interval. It does not do so in those who work in a less stressful environment. The most plausible view is that cognitive failure makes a person vulnerable to showing bad effects of stress, rather than itself resulting from stress.

During the past few years a considerable amount of interest has developed in the assessment of minor everyday slips or errors. This interest is perhaps due more to the work of Reason (1977, 1979) than of any other single individual; starting from the analysis of accident reports, he has progressed to the recording of bizarre episodes of 'absent-mindedness' in diaries kept during everyday life, and to general questionnaires of the frequency with which such episodes are reported by individuals. The lapses which have interested him can be typified by the pilot who sets his course by the wrong end of the compass needle, or who closes down the *surviving* engine when one fails; or who forgets to change fuel tanks. A particularly interesting detailed example is given by Ruffell-Smith (1979), of a flight engineer who repeatedly miscalculated the safe load of fuel to retain in a (simulated) aircraft during an emergency landing, and finally concluded that his instruments must be in error. Similar though less dramatic episodes occur to all of us during the average week; as the examples show, they may involve perceptual failures, or failures of memory, or actions which are misdirected. The common element is that there is a departure from the normal smooth flow of function, and events do not proceed in accordance with intention. The situation is rather different from the failures of *ability* which again we all reach at our limits, and which may arise at quite low levels of difficulty for some individuals. In the lapses we are now discussing, the ability is certainly present and the functions may on other occasions be regarded as laughably simple; but they do sometimes fail.

Quite apart from the initiative of Reason, there are practical and theoretical reasons for the recent interest. Although it is traditional to emphasize affective components of clinical problems, many such problems do also involve disorders of attention, of memory, and of the control of thought or action. In the General Health Questionnaire (Goldberg, 1972), it is interesting to note that the single question most highly discriminating between patients and controls is one relating to failure of concentration. From the armchair, it is tempting to attribute a high frequency of absent-minded errors to stress, though more systematic record keeping does not confirm that people report particular stress at the time of such errors (Reason, 1981). Nevertheless, a small-scale study of laundry workers reported by Broadbent (1980) gave a significantly higher incidence of reported lapses in individuals

whose jobs, as assessed before scrutiny of the lapse data, had a higher number of stressful components such as time pressure or social isolation. In that case, there was a similar correlation between the stress features of the job and poor mental health as assessed by a modified version of the Middlesex Hospital Questionnaire (Crown & Crisp, 1966). As we shall see later, there is much more, and more substantial, evidence that people with a high rate of cognitive failure are also likely to report a high incidence of more conventional affective symptoms. At the least therefore it seems desirable to tease out the relationship between these failures of cognitive control and other symptoms of disorder. For example, there is the possibility that they might furnish an early indication of psychiatric difficulties.

Questionnaires of everyday memory failure have been reported by Herrmann & Neisser (1978) and by Bennett-Levy & Powell (1980); the latter authors make the point that answers to such questions are only weakly correlated with laboratory tests of memory, and this is supported by a case in which actual measurement of successful recall in real life is *negatively* related to success at laboratory recall (Wilkins & Baddeley, 1978). This raises the question of the theoretical interest in this topic. For some years now approaches to human memory have distinguished between the basic structure or facilities of the memory system, and the 'control processes', or strategies, associated with alternative modes of operation of the system. An early example was the seminal work of Atkinson & Shiffrin (1968) on the changes in patterns of recall produced by alternative strategies of rehearsal. The same person, without change in underlying capacities, may forget item A and recall item B if rehearsing in one manner; but recall A and forget B if rehearsing in another. Laboratory tasks tend to induce particular strategies, whereas in real life the problem may well be the choice of the strategy appropriate to the situation. Thus everyday failures may be sensitive to changes which are not revealed in the laboratory, at least with existing techniques.

But the problems of control go further than the topic of memory. Attempts in the field of artificial intelligence to model processes such as speech comprehension or problem solving, encounter as a major problem the sequencing of the various component operations which may be involved. Appropriate analysis of a verbal input may depend upon consultation, at the correct point, of memory for particular past events or for general knowledge of the world; problem solution may depend on the occurrence of a specific action which changes the environment before rather than after possible lines of advance are evaluated (Newell & Simon, 1972; Winograd, 1972; Boden, 1977). If everyday problems of memory are associated with defects in the choice of appropriate strategies for sequencing operations, then we might expect them to be linked with corresponding problems of perceptual selection and of overt action. An analysis of various slips of action and the possible mechanisms underlying them has been presented by Norman (1981).

Hence, the present paper deals with a variety of lapses and not simply with memory failures; it cannot be assumed that perceptual, memory, and action failures will be associated with each other, and one of the major purposes of this inquiry was to see whether they are. The general title, 'cognitive failure', is adopted to cover all types of lapse, whichever of the traditional chapter headings of psychology they occupy.

A second purpose of the study was to examine the relation between cognitive failure and a number of traditional dimensions of individual difference. It could well be urged, for example, that people who report a large number of lapses simply have a poor self-image or are less influenced by social desirability. There are other familiar dimensions of individual difference, such as neuroticism or intelligence; one might, in measuring cognitive failure, be re-inventing such dimensions.

A third purpose, arising out of the second, was to get some assessment of the external validity of people's own reports of cognitive failure. Because of the problem of laboratory

measures, mentioned above, the direct approach was adopted, of comparing self-report with the opinion held about the same person by another who knows them well.

Lastly, some attempt has been made to check whether cognitive failure is indeed a useful dependent variable showing the consequences of stress or early signs of clinical problems. The main line of attack has been to measure stability of scores over time; but in some cases known degrees of stress have been met between test and retest.

Development of the questionnaire

Items

Twenty-five possible failures were assembled from events which had happened to ourselves or acquaintances, the aim being to find episodes which the majority of people would accept as occurring to them at least occasionally. (This goal turned out too optimistic, as will be seen.) The particular events included were also intended to cover each of the three general areas already mentioned. The items chosen, in the wording and with the response format finally adopted, are shown in Appendix 1. It should be noted (see end of paper) that the method of selecting items may well have determined some of the results.

Response format

As can be seen, the respondent is asked to indicate the frequency with which he or she makes such mistakes, the final version being a five-point scale allowing the response of 'never'. Other versions were tried in some of the groups studied; the first used was a six-point scale with specific rates of occurrence attached to the points (once a week, more than once a week, etc.). The advantage of the final version is that since the absolute number of opportunities for error are different for different kinds of event, respondents dislike having to relate them to an absolute scale of time. They tend also to use only four points; but some of them object strongly to having no opportunity to say 'never'. In some of the groups, responses were obtained by trained interviewers who coded the responses into the five-point framework, and the changes in the response categories (as well as the wording finally used for questions) were partly due to the difficulties encountered in these interviews.

Time period

It will also be noted that responses are required with reference specifically to the last six months; again, some respondents were unclear how to reply with earlier versions which had no time period mentioned. The final format also seemed more consistent with the hope that the questionnaire might be sensitive to temporary state rather than underlying traits, and thus useful as an indicator of occupational stresses. In the student nurse groups who responded repeatedly at six-week intervals, responses after the first were given with reference to six weeks rather than months.

Positive/negative wording

All questions are worded in the same direction, rather than adopting the device of wording some questions positively and some negatively to cancel out biases favouring affirmation or denial. A version with reversed wording was tried on a small scale and did not appear to show serious differences in structure or average level of response; in a trial with a mixed wording there was some reason to suspect that respondents were misreading the response scale, and unintentionally answering at the wrong end of the scale. The proper check against performance being due to affirmation bias is rather to examine the correlation with other measures of that bias. It is notable that current practice is for similar reasons swinging against the use of mixed affirmative and negative questions (Goldberg, 1972).

Lie scale questions

Another, and even more serious, possibility is that response to such a questionnaire might be primarily determined by the willingness of the person to admit to damaging facts about themselves. As one check on this, the first version of the CFQ included five extra questions, drawn from the lie scale of the Eysenck Personality Questionnaire. That is, they asked about failings which were social rather than cognitive, such as lying and crossness. The format for response was the same as for the cognitive failures. This of course means that high scores on these questions were in the opposite direction to the usual lie scale, which is high when people are denying failings. Our purpose was to see whether the tendency to admit these social failings was related to the tendency to admit cognitive failure. After the initial trials however these questions were omitted, as they had served their purpose, and as respondents found them objectionable.

Groups studied

The means for the major groups which have been studied, and the main differences in the administration of each, are shown in Table 1. It will be realized that the numbers of variations used mean that normative data are not comparable from one group to another.

Table 1. Some of the main groups examined

Source	Sex	No.	Scale used	Special changes	Mean score (SD in brackets)	MHQ total (SD in brackets)
Oxford Subject Panel	F	98	1-6	—	79.3 (13.32)	21.47 (9.71)
	M	26	1-6	—	79.1 (13.20)	18.31 (8.07)
Cambridge Subject Panel						
Positive wording	F	31	1-4	Pos. wording	51.77 (7.94)	—
Negative wording	F	58	1-4	Neg. wording	52.52 (9.90)	—
NHS laundry workers	F	18	1-4	Interview	44.28 (11.45)	16.39 (8.93)
Management sample	M	171	1-4	—	43.7 (8.40)	13.0 (7.16)
	F	82	1-4	—	44.01 (10.09)	14.23 (7.53)
Student nurses	F	221	0-4	Not restricted to last six months	52.48 (14.52)	14.25 (7.41)
Car factory						
Production workers	M	90	0-4	Interview	35.02 (11.52)	13.03 (7.77)
Skilled men	M	115	0-4	Interview	36.65 (9.41)	11.74 (7.63)

Nevertheless, it was thought this was a lesser evil than continuing with faults in the questionnaire that had already become clear. None of the groups shows results inconsistent with any of the findings given in subsequent sections, although some are more relevant than others under each heading, because of special features of their conditions.

Correlations used

In the analysis, it should be noted that some correlations are product moment, but some are tau rank-order correlations; the latter have been used in studies where doubt was felt about the normality of the distribution of some variable in the study. Great care should be taken in comparing degrees of association in different studies, as tau is numerically smaller than r calculated from the same data.

Lastly we must emphasize the arbitrary nature of significance levels in the results which follow. When very many relationships are examined, some are bound to reach conventional levels.

The common factor of cognitive failure

The first finding is simply that all the questions tend to be positively correlated. Consider the results of the first 98 people, all women members of the Oxford Subject Panel, aged between 20 and 40, responding on the six-point scale, and with lie questions included. Coefficient alpha was 0.79. The correlations between each item and the total of all other

Table 2. Product moment correlations of each individual question with the total *excluding* that question, but including five lie scale questions (data from 98 women of Oxford subject panel)

Q	r	Q	r	Q	r
1	0.239	9	0.258	17	0.455
2	0.372	10	0.328	18	0.354
Lie	0.112	Lie	0.181	19	0.266
3	0.305	11	0.287	20	0.221
4	0.251	12	0.405	21	0.531
5	0.385	Lie	0.215	22	0.466
6	0.256	13	0.255	Lie	0.220
7	0.006	14	0.243	23	0.531
Lie	0.099	15	0.362	24	0.362
8	0.325	16	0.493	25	0.238

items exceed 0.23 for 23 of the questions; the exceptions are the five lie questions, and two others, questions 7 and 20. The lie scale questions are in fact clearly different from the primary questions; each of them has an average of more than eight negative correlations with the 25 primary questions, whereas the latter average only three each out of the 24 possible with each other. Thus it does appear that the social failures are rather different from the cognitive ones. Correspondingly, alpha for the later version of the questionnaire, excluding the lie scale and on a four-point response format, rose to 0.89.

The next question concerns the presence or absence of any further structure amongst the cognitive questions themselves. We have carried out factor analysis and, more especially, multidimensional scaling of the questions; but apart from the obvious general factor the results are highly variable from one group to another, and it seems inadvisable to report any one of them. This is not a matter of rotation of axes, as the distances between questions in the best-fitting space are also variable; the problem is best illustrated by Table 3, which gives the results of intercorrelating the correlation matrices of questions, for a number of groups. As can be seen, the values obtained are quite low, meaning that the questions most similar in one group of people have only a slight tendency to be the most similar in another;

one cannot then expect stable factors or dimensions. Three deviant questions do indeed seem to form a cluster in the majority of groups and by most forms of analysis; these are the questions 7 and 20 already mentioned as showing a low item – total correlation, together with question 11. For example questions 7 and 20 correlate 0.359 with each other but with nothing else higher than 0.28; while question 11 correlates 0.421 with a lie scale question but no higher than 0.28 with any other question. These three questions involve one perceptual and one memory failure, each concerning people's names, and one action failure (not answering letters). Conceivably they have a social rather than a purely cognitive component.

Table 3. The result of taking the matrix of RHO correlations between questions in individual groups, and then calculating a product moment correlation between the matrices for each pair of groups (only one correlation is over 0.3; the structure is clearly highly variable)

	F98	M26	MH	HF	NAP	POS
Male subject panel	M26 0.109	—	—	—	—	—
Male managers	HM 0.281**	0.230**	—	—	—	—
Female from management group	HF 0.105	-0.011	0.198**	—	—	—
Laundry workers	NAP 0.085	0.159**	0.128*	0.049*	—	—
Cambridge subject panel	POS 0.149**	0.111	0.065	-0.046	-0.012	—
with wording Pos. or Neg.	NEG 0.191**	0.207**	0.315**	0.244**	0.148*	0.088

* $P < 0.05$; ** $P < 0.01$.

However, even these questions still have a positive correlation with the total. Without prejudice to any future splitting of the types of cognitive failure into subdivisions, it seems advisable for the present to look at the properties of the total score for the whole 25 questions, and the remainder of this paper confines itself to the relationships of this total score. It should also be emphasized that there is no evidence for separate categories of perceptual, memory, and action failure.

Further evidence for the generality of the tendency to cognitive failure comes from data most kindly provided by Reason. He has administered the CFQ to 94 psychology undergraduates, who had also completed the Short Inventory of Memory Experiences of Herrmann & Neisser, Reason's own Absent-mindedness Questionnaire, and two forms of his Slips of Action Inventory; the correlations are given in Table 4. It can be seen that all the inventories are fairly closely correlated, whichever aspect of cognitive failure they attempt to measure.

Table 4. Product moment correlations between CFQ, the Hermann & Neisser memory questionnaire, and Reason's inventories of slips of action and of absent-mindedness (data on 94 undergraduates, courtesy of Professor Reason)

	Slips of action Form A	Slips of action Form B	Absent-mindedness questionnaire	Short inventory of memory experiences (forgetting scale)
CFQ	0.57	0.58	0.62	0.59
SAI(A)	—	0.81	0.64	0.53
SAI(B)	—	—	0.61	0.52
AMQ	—	—	—	0.58

Relative independence of other measures

A general factor in the responses need not of course imply that all types of failure are due to defective control strategies in the person. An obvious and less romantic explanation would be a general willingness to admit defects, of the kind already suggested. The discrepancy between the lie scale questions and the remainder, which we saw in the last section, argues against this; but further evidence is necessary. Table 5 shows the correlation between CFQ and a number of more familiar measures, obtained from a sample of student nurses as part of a study of occupational stress. A full report of this study is still in preparation (Parkes,

Table 5. Product moment correlation of CFQ and some standard psychometric measures, for 221 student nurses

	EPQ E	EPQ N	EPQ P	EPQ L	Rotter external control	Spielberger trait-anxiety	ACL Defensiveness
CFQ	-0.113	0.277**	0.070	-0.177**	0.350**	0.311**	-0.246**
E	—	-0.201**	0.080	-0.235**	-0.092	-0.400**	0.130
N	—	—	-0.072	-0.078	0.303**	0.687**	-0.197**
P	—	—	—	-0.177*	0.082	0.016	-0.327**
L	—	—	—	—	-0.130	0.012	0.223**
Ext. control	—	—	—	—	—	0.300**	-0.084
Trait-anxiety ¹	—	—	—	—	—	—	-0.300**

* $P < 0.05$; ** $P < 0.01$.

1981) although a preliminary account has appeared (Parkes, 1980a). Briefly, 221 nurses were assessed before starting ward duties and various base-line measures taken. The closest relationship of the CFQ is to Rotter's measure of externality of control (Rotter, 1966); other correlations are numerically small. It will be noted that, on this fairly large sample, CFQ does correlate significantly both with the N scale of the EPQ, and (negatively) with the L scale and with the defensiveness measure from the Gough Adjective Checklist (Gough & Heilbrun, 1965). Equally however the L scale is significantly negatively related to extraversion and to the P scale, defensiveness to the N and P scales, and extraversion to the N scale. All these relationships while real are very weak, amounting to 6 or 7 per cent of the variance. They contrast with the massive correlation of neuroticism with trait anxiety (Spielberger *et al.*, 1970) amounting to 45 per cent of the variance and with the relationships of the CFQ to other measures of cognitive failure, shown in Table 4, and accounting for about 35 per cent of the variance. It seems fair to conclude that contamination of CFQ by factors of defensiveness or social desirability is no greater than that of the other traditional dimensions. In each case the contamination is small compared with the major distinct quantity being measured.

It should be pointed out at this stage that social desirability of the type measured by the L scale is distinguishable from a general tendency to denial; the former could be described as the tendency to give a response appropriate to the situation, while the latter would be more the tendency to respond in such a way as to preserve self-esteem (Parkes, 1980b). The CFQ may well be affected by both factors, as several of the other measures are, but at least it appears to be scoring something separate from them.

In the same way, the CFQ does not appear to be very closely related to test intelligence or to educational level; data from an Edinburgh sample kindly supplied by David Weeks give a product moment correlation of -0.157 with the Mill Hill ($n = 128$); and of -0.15

($n = 51$) with the Progressive Matrices. Similarly, none of our samples from the Oxford Subject Panel (which has of course a wider range than our occupational samples) has shown any significant correlation with educational level. It may still be of course that larger-scale studies establish some relationship; but if so it must certainly be small.

External criterion

A lack of relationship to other tests is promising, but could be achieved by the trivial mechanism of respondents answering high or low according to the whim of the moment. Some external measure of the same people is needed. We were fortunate in obtaining the collaboration of Dr Andrew Melhuish, whose longitudinal study of stress in middle management has secured responses from a number of such managers and also from their wives (see Cooper & Melhuish, 1980). Each partner was asked to complete the CFQ for themselves, and also to assess their partner on a number of rating scales; the latter are given as Appendix 2 — the CFQ-for-others. The scales are intended to pick up those aspects of cognitive failure which are apparent to others; unfortunately some failures are likely to remain unknown to anybody except the person who suffers them, so we could not ask the wife to predict the actual CFQ responses of the husband, or vice versa. In each case, the ratings and CFQ were completed and handed back before any opportunity to compare notes with the partner.

In general, there was only a very weak relationship between the CFQ scores of the husband and wife; even including a few extra couples who had not done the CFQ-for-others, to bring n up to 67, the tau correlation was still insignificant at 0.144. The ratings of one by the other using the CFQ-for-others had a slightly greater tendency to similarity, helped by some relationships which seemed to reflect the degree of mutual approval. For instance, there are significant correlations showing that the wife is likely to rate the husband as unduly cross if he rates her as absent-minded and clumsy. The total of all eight ratings by the husband is therefore just significantly correlated with those by the wife, $\tau = 0.243$. Table 6 shows however that the total rating on the CFQ-for-others gives comfortably higher correlations with the partner's CFQ, both for the man's rating of the woman and for the woman's rating of the man. Slightly different ratings seem to be responsible in the two cases, the husband's rating of his wife's inability to make up her mind showing little relation

Table 6. Tau correlations between ratings of cognitive failure by marital partner, and self-reported CFQ and MHQ

Partner's rating	Husband's CFQ	Wife's CFQ	Husband's MHQ	Wife's MHQ
Absent-mindedness	0.247*	0.324**	0.195	0.440**
Lack of concentration	0.209	0.173	0.071	0.329*
Forgetful	0.316**	0.280*	0.144	0.239*
Failure to notice	0.345**	0.197	0.147	0.281*
Clumsy	0.236	0.230	-0.142	0.230
Difficulty in making up mind	0.324**	0.161	-0.072	0.330*
Disorganized	0.118	0.243	-0.060	0.580**
Cross	0.015	0.200	0.059	0.337**
Totals of ratings	0.315**	0.360**	0.069	0.486**
Mean of questionnaire scores (SD in brackets)	43.32 (9.38)	46.17 (11.30)	14.41 (8.02)	14.92 (8.09)

* $P < 0.05$; ** $P < 0.01$.

to her CFQ, while her rating of the same type gives a highly significant relationship to his. In general however the key point is that the respondent's own view of his or her liability to cognitive failure does tend to be shared by somebody who has a good opportunity to judge.

In the light of certain later findings, it is worth drawing attention here to another point in Table 6. Both husbands and wives had completed the modified version of the Middlesex Hospital Questionnaire, already mentioned in connection with laundry workers. The MHQ is a symptom inventory which discriminates neurotic patients from others (Crown & Crisp, 1966; Crisp *et al.*, 1978). It is however intended more as an assessment of state than of underlying personality as is, say, the EPQ. The husband's MHQ score showed no significant relation to his wife's rating of his cognitive failures; but his rating of her, for seven of the rating scales, actually correlated more highly with her MHQ than with her CFQ. What he is rating corresponds more to her self-report of emotional than of cognitive state; what she is rating corresponds more to his self-report of cognitive than of emotional state.

Two suggestive findings from other studies may be mentioned here. First, a sample of people who came forward in response to an appeal for former ECT patients gave higher scores than a normal control group (Freeman *et al.*, 1980). Second, a small number of the student nurse sample (12) have got as far as final-year examinations, but failed. The CFQ of these nurses was higher than that of the 24 who passed at the same exam; the significance was however dubious, $t = 1.79$.

Both these findings show some external event corresponding to the person's own opinion. However, neither is totally objective. The inclusion of people in the ECT sample reflects their own opinions. A follow-up of depressed patients, from a time before allocation to ECT or other treatments, did not show corresponding deficits associated with ECT (Weeks *et al.*, 1980). Similarly, the marking of nurses' examination papers was undertaken by people who knew them. Thus these findings cannot be taken as proving more than the study of married couples has done; self-report on the CFQ has an external importance, and is not a whim of the moment.

Stability of scores

Thus far, all has gone well for the notion that cognitive failure is an indicator of general disturbance of control, and perhaps a warning of stress or of other problems. The various kinds of failure seem to occur in the same person rather than separating into perceptual, memory, and motor failures; the tendency to report them is not equivalent to neuroticism, social desirability set, or low intelligence; and there is some evidence that high scores do correspond to a true liability to make such failures rather than simply to report them. However, if cognitive failures reflected temporary state the score of any individual should change from time to time, as circumstances change. It should not remain constant as we expect a trait to do. In this section we see evidence that the scores are in fact rather stable.

There are three cases in which people have replied to the questionnaire at least twice over long intervals. First, David Weeks in Edinburgh has again contributed results for two groups, one with retesting occurring over an average interval of 21 weeks, and one with retesting over an average interval of 65 weeks. For the first group, the test-retest correlation gave $r = 0.824$ ($n = 57$); for the second group, $r = 0.803$ ($n = 32$). In each group, it was also notable that there was no correlation between the size for an individual of the test-retest interval and of the difference in score.

Second, the managers studied by Andrew Melhuish have in a number of cases returned questionnaires a second time after an interval of between one and two years; for the 114 who could be compared in this way, $\tau = 0.621$. Remembering the fact that tau gives values

numerically lower than r , this agrees quite well. For a smaller group of wives ($n = 42$) who had also been followed over the same period, $\tau = 0.563$.

Third, the student nurses of Parkes (1980a) were tested repeatedly at six-week intervals throughout the period of their training. There were several factors which might well lower test-retest correlations in this group; as one might expect, they were young, undergoing numerous life-events, and were known to have been exposed in their work to conditions which (for example) gave significant changes in their scores on the modified MHQ. Furthermore, in the case of the CFQ, trials after the first used a different format from the first by asking how frequently the failure had occurred in the last six weeks. This was of course necessary to make the repeated tests meaningful, but would lower initial-final correlations.

For 73 nurses who have so far been observed for 16 months, the initial-final product moment correlation was 0.54. On the same individuals, the initial-final correlations were 0.61 for N, 0.72 for E, 0.55 for P, and 0.64 for L. Spielberger's Trait Anxiety gave 0.65, and Rotter's Internal-External Control 0.69; and the MHQ (for $n = 82$) 0.50. Thus the CFQ is nearly as stable as those more familiar measures, which are usually regarded as indices of trait rather than of temporary state.

These results would have been encouraging if we had been seeking a persistent characteristic of the individual; they do not however hold out much hope that the CFQ will show up temporary disruption of the individual.

Relation to psychiatric symptoms

Mention has already been made of the MHQ, which calls for self-report of mild symptoms under a number of headings; a modified set of questions based upon it have been given to a number of our groups. The modifications were mostly small changes of wording to make the questionnaire acceptable for industrial use, the elimination of the phobic and hysteria scales for similar reasons, and in particular the change of all possible questions to refer to the last six weeks. That is, our modification is very much a measure of state rather than trait. Correlations from a number of groups are given in Table 7.

Table 7. Tau correlations of CFQ and scales from the MHQ, for various groups

	No.	Anxiety	Somatic	Depression	Obsessional		Total
					Symptoms	Personality	
Male subject panel	26	0.314*	0.293*	0.366*	0.345*	0.133	0.412**
Female subject panel	34	0.352*	0.396**	0.268*	0.196	-0.017	0.414**
Management group	171	0.275**	0.218**	0.291**	0.236**	0.062	0.271**
Management wives	78	0.252**	0.124	0.296**	0.290**	0.021	0.262**
Female laundry workers	18	0.408*	0.396*	0.450*	(0.378)*		0.541**
Student nurses	101	0.261**	0.158*	0.191**	0.261**	-0.101	0.221**

* $P < 0.05$; ** $P < 0.01$.

It will be seen that in each case the correlation of total CFQ and total MHQ is significant beyond the 0.01 level, and similarly the individual correlations with the anxiety and depression scales are all significant. (In comparing this table with, say, Table 5 remember that one table refers to tau and the other to product moment correlations. As a guide to the relative magnitudes, we constructed three samples of artificial data for each of which tau was 0.2; the average value of r , calculated from the same data, was 0.395. For three samples

for which tau was 0.3, the average value of r was 0.58. Thus even the smallest correlation of CFQ with total MHQ, in Table 7, represents a closer degree of relationship than the largest correlation of CFQ with any other measure in Table 5.) Indeed, the only negative correlations with the CFQ occur for the 'obsessional personality' scale, which consists of certain questions from the original MHQ obsessiveness scale that are now known to separate from the remainder on factorial analysis (Crisp *et al.*, 1978) and which in general do not seem to correlate with other MHQ scales. The results in Table 7 are typical rather than exhaustive of findings in other occupational groups we have examined; thus it does appear that people with a high rate of cognitive failure are also people with a large number of minor symptoms of other kinds.

Yet, we have seen that CFQ tends to be more stable than might be expected for a measure of state; and we may recall the finding of Reason (1981) that people keeping records of cognitive failures say that they do not occur particularly often when stress is experienced. What, then, produced the correlation?

The study of student nurses by Parkes (1980a, 1981) sheds considerable light on this point; because it is unusual in that the nurses were given the CFQ before entering into the stressful situation of working on hospital wards. It was found that wards differed in the degree of difficulty which they produced for the nurse; for example, a sample of 88 nurses who had each worked on medical and on surgical wards gave significantly different MHQ scores when they had spent the previous six weeks on medical wards than when they had been working on surgical wards. It is relevant that their CFQ did *not* change significantly.

The precise reasons for regarding a ward as stressful may of course differ quite widely, but broadly one may follow Karasek (1979, 1980) in distinguishing two important characteristics. One is the quantity of work demanded and the other is the degree of freedom or autonomy available to the individual in meeting the demands. The latter factor reduces the effect of the former. Each ward was rated by the nurses on these two dimensions, using the questions found by Karasek to be relatively pure measures of each in large-scale social surveys of workers in Sweden and the USA.

The rating given by a nurse might conceivably be contaminated by her own state of health, much as people living in the same district may report more annoyance from the same level of aircraft noise if their own level of mental health is poor (Tarnapolsky & Morton-Williams, 1980). For each nurse therefore the mean difference between the two dimensions was taken for the ratings given by all *other* nurses who had worked on the same ward as the one where the individual spent her first 12 weeks. Again, CFQ showed no significant change in high and low stress wards as assessed in this way. However, the nurses could be divided into two groups, $n = 48$ and $n = 53$, who had worked on relatively high stress wards and on relatively low stress wards as their first assignment after their CFQ had been determined.

In the low stress wards, previous CFQ and subsequent MHQ (total of two determinations) were effectively uncorrelated, $\tau = 0.087$; whereas in the high stress wards the two scores gave $\tau = 0.462$ which was highly significant. Thus it seems that the level of minor symptoms after six weeks of work is related to the earlier CFQ only if the work is stressful.

High CFQ, in other words, is a vulnerability factor making the individual less able to resist the effects of stress. It is tempting to recall the views of Lazarus (1966), that the initial impact of a stressful situation is to produce an attempt at cognitive appraisal and from this a strategy of coping can be developed. Only if coping is unsuccessful does the person develop symptoms.

A minor line of evidence pointing in the same direction should be mentioned at this point. One sample studied was a series of 51 neurotic patients at the Warneford Hospital, who were also given a Standard Psychiatric Interview (Goldberg *et al.*, 1970). Some

difficulty was obtained in getting fair responses to a number of CFQ questions, since some of the patients were not in the previous six weeks exposed to some of the situations (e.g. shopping, driving). After excluding such cases as indeterminable, 34 cases remained, and for these the mean CFQ was 44.88. Reference to Table 1 will show that this is well within the range met in other groups, although the MHQ of this sample was 29.83 and thus twice as high as the majority of groups. On the SPI, different results were obtained in the male ($n = 20$) and female ($n = 14$) subgroups. In the former, CFQ correlated with the psychiatrist's assessment of obsessional symptoms, $\tau = 0.362$, while in the latter there were correlations with fatigue and with lack of concentration, 0.629 and 0.535 respectively. No other relationships were significant. These results again do not support a view that cognitive failures appear as a result of, say, high anxiety. It is therefore quite hard to regard CFQ as a response to stress; the repeated correlations with MHQ are better interpreted by seeing CFQ as a vulnerability factor

Objective measures

If CFQ measures some fairly enduring trait, then more than ever one would wish to find some test more objective than a self-report scale for assessing the level of cognitive failure in an individual. By such means one might hope to escape the problem of defensive unwillingness to admit error, already mentioned. The difficulty is to identify the right objective function to test. Thus far, we have found no significant relationship on groups of $n = 17-20$, between CFQ and immediate memory for nine-item lists visually or acoustically presented, for the size of effect of acoustic suffix, for the size of the effect of articulatory suppression, for efficiency at identification of blurred words, for longer-term memory for categorized material; and for performance on the Williams delayed recall test. (The latter has been found to give a significant effect by David Weeks in Edinburgh, on a group including depressives, and it may be that a larger sample might give a significant result.) Perhaps the most disappointing negative finding concerns the combination of two tasks. In a school project, a team of young investigators had found that children who reported many cognitive failures in a diary-keeping technique were also inferior at combining two tasks; specifically, counting backwards and tracing a star in a mirror (Bexhill, 1980). We have attempted to show a similar relationship with adults, in two slightly different replications of their study, but have failed to get any relationship to the CFQ. It may be that the reduced variance due to age in a school sample may have helped their result, as we found that older people showed very large deteriorations in ability to combine tasks. Alternatively, it may be that the special stress of an investigation carried out by people known to the subjects might create a relationship which disappears in the less involved situation of the laboratory. In any event, it is clear that CFQ does not relate so closely to inability to combine two tasks that it will show up reliably in groups of 20 or so. The same must also be said of the other functions we have mentioned.

It is however still worth some further effort to seek for an objective correlate of high CFQ. Harris & Wilkins (1981) have found one weak but significant relationship with an objective measure of performance; they required subjects to watch a film, about which they would later be questioned. Behind them was a clock, which they could consult at will. Each subject had a series of target times, spaced at three- or nine-minute intervals, and was asked to hold up a sheet to a TV camera as each target time arrived. They were more than 15 seconds late on rather more than 10 per cent of trials. The number of occasions on which an individual was late was greater in high CFQ individuals, $P < 0.05$, on 27 subjects. This task is manifestly rather different from most laboratory measures of performance and it may well be therefore that further developments of such an approach will prove fruitful.

Conclusions

The CFQ appears therefore to be promising, but in a way rather different from that originally envisaged. It does seem to measure a general liability to failure, which is distinguishable from the tendency to admit discreditable things or from neuroticism; and which has an importance in the external world because the opinion of the persons themselves is shared by others who know them. From a theoretical point of view, it is quite important that memory, action, and perception are all reported as affected in the same people, which supports the notion of some defect of overall control. But, the hope that CFQ could be used to show changes of state resulting from stress has so far been unfulfilled. This may well be due to the method of assembling items, which probably selected those failures that occur repeatedly over a long period of time. They would thus be unsatisfactory as measures of state. Further efforts to construct some different measure of state might be worthwhile.

On the other hand, we have much better evidence that high CFQ increases vulnerability to subsequent stress. This is reasonable, in terms of the view that stress has its major effects on those who cannot cope cognitively. Yet it leaves obscure the origins of the liability to cognitive failure itself; and we still have no satisfactory method of measuring that liability, other than self-report. At least the results assembled so far make the topic worth further investigation.

Acknowledgements

As has been noted in the text, very useful data, and also comments, have been supplied by Andrew Melhuish of Henley, David Weeks of Edinburgh, Arnold Wilkins of the Applied Psychology Unit, Cambridge, and Jim Reason of Manchester; we are most grateful. Chris Robertson has also helped enormously with statistical advice and by conducting a number of alternative analyses of the data. The study of student nurses is supported by the Health and Safety Executive; the remainder of the research was supported by the Medical Research Council.

References

- Atkinson, R. C. & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (eds), *Psychology of Learning and Motivation*, Vol. 2. New York: Academic Press.
- Bennett-Levy, J. & Powell, G. E. (1980). The Subjective Memory Questionnaire (SMQ). An investigation into the self-reporting of 'real-life' memory skills. *British Journal of Social and Clinical Psychology*, **19**, 177-188.
- Bexhill VI Form College (1980). An investigation into the incidence and causes of absent-minded behaviour. Report to the BBC Young Scientist of the Year Competition.
- Boden, M. (1977). *Artificial Intelligence and Natural Man*. Brighton: Harvester.
- Broadbent, D. E. (1980). Chronic effects from the physical nature of work. In B. Gardell & G. Johansson (eds), *Working Life: A Social Science Contribution to Work Reform*. London: Wiley.
- Cooper, C. L. & Melhuish, A. (1980). Occupational stress and managers. *Journal of Occupational Medicine*, **22**, 588-592.
- Crisp, A. H., Gaynor-Jones, M. & Slater, P. (1978). The Middlesex Hospital Questionnaire: A validity study. *British Journal of Medical Psychology*, **51**, 269-280.
- Crown, S. & Crisp, A. H. (1966). A short clinical diagnostic self-rating scale for psycho-neurotic patients. *British Journal of Psychiatry*, **112**, 917-923.
- Freeman, C. P. L., Weeks, D. & Kendell, R. E. (1980). ECT II: Patients who complain. *British Journal of Psychiatry*, **137**, 17-25.
- Goldberg, D. P. (1972). *The Detection of Psychiatric Illness by Questionnaire*. London: Oxford University Press.
- Goldberg, D. P., Cooper, B., Eastwood, M. R., Kedward, H. G. & Shepherd, M. (1970). A standardized psychiatric interview for use in community surveys. *British Journal of Preventative and Social Medicine*, **24**, 18-23.

- Gough, H. G. & Heilbrun, A. B. (1965). *The Adjective Checklist Manual*. Palo Alto, CA: Consulting Psychologists Press.
- Harris, J. E. & Wilkins, A. J. (1981). Remembering to do things: A theoretical framework and an illustrative experiment. Submitted for publication.
- Herrmann, D. & Neisser, U. (1978). An inventory of everyday memory experiences. In M. M. Gruneberg, P. E. Morris & R. N. Sykes (eds), *Practical Aspects of Memory*. London: Academic Press.
- Karasek, R. A. (1979). Job demands, job decision latitude, and mental strain: Implications for job redesign. *Administrative Science Quarterly*, **24**, 285–308.
- Karasek, R. A. (1980). Job socialisation and job strain. The implications of two related psychosocial mechanisms for job design. In B. Gardell & G. Johansson (eds), *Man and Working Life*. London: Wiley.
- Lazarus, R. S. (1966). *Psychological Stress and the Coping Process*. New York: McGraw-Hill.
- Newell, A. & Simon, H. A. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Norman, D. A. (1981). Categorization of action slips. *Psychological Review*, **88**, 1–15.
- Parkes, K. R. (1980a). Occupational stress among student nurses: 1. *Nursing Times*, **76**, 113–116; Occasional Paper no. 25.
- Parkes, K. R. (1980b). Social desirability, defensiveness, and self-report psychiatric inventory scores. *Psychological Medicine*, **10**, 735–742.
- Parkes, K. R. (1981). Occupational stress among student nurses: A natural experiment. In preparation.
- Reason, J. T. (1977). Skill and error in everyday life. In M. Howe (ed.), *Adult Learning*. London: Wiley.
- Reason, J. T. (1979). Actions not as planned: The price of automatisisation. In G. Underwood & R. Stevens (eds), *Aspects of Consciousness*. London: Academic Press.
- Reason, J. T. (1981). Lapses of attention. In R. Parasuraman, R. Davies & J. Beatty (eds), *Varieties of Attention*. New York: Academic Press.
- Rotter, J. B. (1966). Generalised expectancies for internal versus external control of reinforcement. *Psychological Monograph 80*: Whole no. 609.
- Ruffell-Smith, H. P. (1979). A simulator study of the interaction of pilot workload with errors, vigilance, and decisions. *NASA Technical Memo 78482*, Scientific and Technical Information Office.
- Spielberger, C. D., Gorsuch, R. L. & Lushene, R. E. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Tarnapolsky, A. & Morton-Williams, J. (1980). *Aircraft Noise and Prevalence of Psychiatric Disorders*. London: Social and Community Planning Research.
- Weeks, D., Freeman, C. P. L. & Kendell, R. E. (1980). ECT: III: Enduring cognitive deficits? *British Journal of Psychiatry*, **137**, 26–37.
- Wilkins, A. J. & Baddeley, A. D. (1978). Remembering to recall in everyday life: An approach to absent-mindedness. In M. M. Gruneberg, P. E. Morris & R. N. Sykes (eds), *Practical Aspects of Memory*. London: Academic Press.
- Winograd, T. (1972). *Understanding Natural Language*. Edinburgh: Edinburgh University Press.

Received 16 February 1981; revised version received 6 May 1981

Requests for reprints should be addressed to Dr Donald E. Broadbent, Department of Experimental Psychology, South Parks Road, Oxford OX1 3UD, UK.
The co-authors are also at the University of Oxford.

Appendix 1

The following questions are about minor mistakes which everyone makes from time to time, but some of which happen more often than others. We want to know how often these things have happened to you in the last six months. Please circle the appropriate number.

	Very often	Quite often	Occasionally	Very rarely	Never
1. Do you read something and find you haven't been thinking about it and must read it again?	4	3	2	1	0
2. Do you find you forget why you went from one part of the house to the other?	4	3	2	1	0
3. Do you fail to notice signposts on the road?	4	3	2	1	0
4. Do you find you confuse right and left when giving directions?	4	3	2	1	0
5. Do you bump into people?	4	3	2	1	0
6. Do you find you forget whether you've turned off a light or a fire or locked the door?	4	3	2	1	0
7. Do you fail to listen to people's names when you are meeting them?	4	3	2	1	0
8. Do you say something and realize afterwards that it might be taken as insulting?	4	3	2	1	0
9. Do you fail to hear people speaking to you when you are doing something else?	4	3	2	1	0
10. Do you lose your temper and regret it?	4	3	2	1	0
11. Do you leave important letters unanswered for days?	4	3	2	1	0
12. Do you find you forget which way to turn on a road you know well but rarely use?	4	3	2	1	0
13. Do you fail to see what you want in a supermarket (although it's there)?	4	3	2	1	0
14. Do you find yourself suddenly wondering whether you've used a word correctly?	4	3	2	1	0
15. Do you have trouble making up your mind?	4	3	2	1	0
16. Do you find you forget appointments?	4	3	2	1	0
17. Do you forget where you put something like a newspaper or a book?	4	3	2	1	0
18. Do you find you accidentally throw away the thing you want and keep what you meant to throw away – as in the example of throwing away the matchbox and putting the used match in your pocket?	4	3	2	1	0
19. Do you daydream when you ought to be listening to something?	4	3	2	1	0
20. Do you find you forget people's names?	4	3	2	1	0
21. Do you start doing one thing at home and get distracted into doing something else (unintentionally)?	4	3	2	1	0
22. Do you find you can't quite remember something although it's 'on the tip of your tongue'?	4	3	2	1	0
23. Do you find you forget what you came to the shops to buy?	4	3	2	1	0
24. Do you drop things?	4	3	2	1	0
25. Do you find you can't think of anything to say?	4	3	2	1	0

Appendix 2: CFQ-for-others

The questions given below are about mistakes and difficulties which everybody has from time to time. We want to know how often, in your opinion, your relative or partner has shown any of these troubles during the last six months. After each question please tick only one of the five possible answers. Please make sure you read them carefully because for some questions 'very often' is on the left side of the page and 'never' is on the right, but for other questions 'never' is on the left and 'very often' is on the right.

During the last six months has your relative or partner seemed to be:

1. Absent-minded, that is making mistakes in what he/she is doing because he/she is thinking of something else?

Very often	Quite often	Occasionally	Very rarely	Never
------------	-------------	--------------	-------------	-------
2. Finding it difficult to concentrate on anything because his/her attention tends to wander from one thing to another?

Never	Very rarely	Occasionally	Quite often	Very often
-------	-------------	--------------	-------------	------------
3. Forgetful, such as forgetting where he/she has put things, or about appointments, or about what he/she has done?

Very often	Quite often	Occasionally	Very rarely	Never
------------	-------------	--------------	-------------	-------
4. Busy thinking about his/her own affairs and so not noticing what is going on around him/her?

Never	Very rarely	Occasionally	Quite often	Very often
-------	-------------	--------------	-------------	------------
5. Clumsy, for example, dropping things or bumping into people?

Very often	Quite often	Occasionally	Very rarely	Never
------------	-------------	--------------	-------------	-------
6. Having difficulty in making up his/her mind?

Never	Very rarely	Occasionally	Quite often	Very often
-------	-------------	--------------	-------------	------------
7. Disorganized, that is, getting into a muddle when doing something because of lack of planning or concentration?

Very often	Quite often	Occasionally	Very rarely	Never
------------	-------------	--------------	-------------	-------
8. Getting unduly cross about minor matters?

Never	Very rarely	Occasionally	Quite often	Very often
-------	-------------	--------------	-------------	------------