Method of test administration as a factor in test validity

The use of a personality questionnaire in the prediction of cancer and coronary heart disease

Grossarth-Maticek, R.; Eysenck, H. J.; Boyle, G. J.

Method of test administration as a factor in test validity: the use of a personality questionnaire in the prediction of cancer and coronary heart disease

R. GROSSARTH-MATICEK, H. J. EYSENCK and G. J. BOYLE

INTRODUCTION

The validity of personality inventories has always presented considerable problems, as we are usually dealing with hypothetical constructs, rather than intervening variables (Cronbach and Meehl, 1955). Intervening variables are directly observable and reducible to empirical laws, while hypothetical constructs refer to processes or entities that are not directly observed (Garber & Strassberg, 1991). A construct, according to Cronbach and Meehl (1955), is defined as "'some postulated attribute of people, assumed to be reflected in test performance" (p. 283). The difficulty with such constructs is that they are basically unobservable and can only be measured indirectly. "Their existence is inferred through the relations between variables that can be observed and the performance on the tests that presumably measure the hypothesized entity" (Garber & Strassberg, 1991, p. 220). Examples of hypothetical constructs are traits, types, attributes, or any quality that cannot easily be operationally defined.

There are four types of validity recognized by Cronbach and Meehl (1955), namely content, concurrent, predictive and construct validity; of those, Loewinger (1957) argued, the first three are essentially ad hoc, and "'construct validity is the whole of validity from a scientific point of view" (p. 636). However that may be, construct validity is necessary "where verifying operations against which to check tests are not automatically available" (Campbell, 1960, p. 550). Construct validity implies the existence of a *nomological network* which is defined as an interlocking system of lawful relations that comprise a theory, and this network consists of both theoretical constructs and observable properties or operations. It is assumed that the theoretical constructs are expressed fully in the questionnaire used to measure the construct in question, and that validation consists in deriving deductions from the theory which are empirically testable. This simple relationship may require some important qualifications. We have chosen for discussion and experimental investigation the concepts of cancer-prone and coronary heart disease (CHD)-prone personalities (Eysenck, 1991), which have become the focus of much interest and numerous empirical investigations in recent years.

Briefly, theory states that some people are more likely than others to contract and die of cancer, while others are more likely to die of CHD. The theory of "Type A personality" as developed by Friedman and Rosenman characterizes the CHD-prone person in such terms as aggressive, hostile, angry, as contrasted with the more normal, healthy "Type B personality" (Eysenck, 1990a). The cancer-prone type (Type C--Temoshok & Dreher, 1992) has been defined in terms of such features as suppression of emotions and inability to cope with stress, leading to feelings of hopelessness, helplessness, and finally, depression (Eysenck, 1991). Questionnaires to measure these hypothetical personality types have been constructed (e.g. Jenkins, Zyzanski, Ryan, Flessas

& Tannenbaum, 1977; Eysenck & Fulker, 1983, for CHD; Grossarth-Maticek & Eysenck, 1990, for cancer). There are two ways of providing construct validity, namely (1) demonstrating that probands high on Type C are diagnosed as suffering from *cancer*, while probands high on Type A are diagnosed as suffering from *coronary heart disease*, more frequently than vice versa, and that probands of Type B are found to be healthy. This method has often been used successfully, but of course suffers from the obvious criticism that the questionnaire response may be in part a function of the illness in question.

(2) More convincing are studies involving healthy probands, tested and followed up for periods of 10 years or more: theory predicts that probands of Type A should die preponderantly of CHD, probands of Type C of cancer, with probands of Type B surviving longer than Type A or Type C probands. There is a good deal of evidence that large-scale findings support the theoretical predictions (Eysenck, 1991; Temoshok & Dreher, 1992). However, the usual argument implies that predictive qualities are inherent in the questionnaire, irrespective of method of administration, and the inevitable subjective reactions of the testee to the whole situation of being asked intimate questions, the purpose of which may not be at all clear to him or her. Thus Amelang and Schrnidt-Rathjens (1992) have published a study using a modified form of the Grossarth-Maticek and Eysenck (1990) inventory, but using a method of simply handing out the questionnaires to probands, whereas in the Grossarth-Maticek studies questionnaires were administered by carefully trained interviewers. Amelang and Sehmidt-Rathjens (1992) did find that sufferers from CHD and cancer had higher scores on Type A and Type C inventories, but failed to discover differences between cancer and CHD patients. This failure may be due to differences in method of administration, and in a previous study, Grossarth-Maticek, Eysenck and Barrett (1993) have shown that this is the most likely explanation. In this first study we have tested the hypothesis that different methods of administering a questionnaire produce differential approximations to truthful admission of undesirable personality traits and behaviours. Four different methods of administration produced different levels of trust and understanding, using the current prediction for healthy *S*s of death by cancer or CHD 13 years later as the criterion. There were significant differences in the accuracy of the predictions, depending crucially on the method of administration of the questionnaires. Best predictions were achieved for Ss when both trust and understanding had been increased by interviewers' suitable participation: worst results were achieved for *S*s when no special effort was made to increase either. Intermediate results were found for procedures which increased either trust or understanding. It is argued that the success or failure of studies investigating the influence of personality and stress on diseases like cancer and coronary heart disease may depend crucially on the adopted method of interrogation.

This conclusion may be of general importance in psychometric personality testing, and it seems appropriate to test our conclusion again in a study based on an altogether larger population. It is very possible that results may depend to some extent on the nature of the questionnaire, and the complexity of the questions. The more intimate the question, the greater the need for trust. The more complex the questions, the greater the need for understanding their meaning. The Grossarth-Maticek questions are both very intimate and very complex, in order to do justice to a theory that is inherently complex; hence they may be particularly in need of interviewer administration. But a similar relationship has been found for the Friedman-Rosenman Type A Type B inventories, where structured interview results have been far more predictive than questionnaire results (Eysenck, 1990a). Personal interaction seems likely to elicit more honest answers, although simple questions dealing with matters not too closely touching upon

the testee's emotional problems may not be subject to this effect to the same extent. It is worth noting, however, that even simple questions concerning smoking habits elicit a sufficiently large number of factually inaccurate answers to make the results useless for the study of the alleged 'passive smoking' effects on health (Lee, 1988).

The aim of the study, then, is to provide additional information on the effects of inventory administration on response accuracy, using as our criterion the relative success of prediction of death from cancer or coronary heart disease. We shall also report on the possibility of improving the questionnaires used in such a way that they can be used without such interviewer assistance. Clearly interviewer administration is costly and time-consuming, and if it can be avoided so much the better.

EXPERIMENT 1

The questionnaire used has been described elsewhere (Grossarth-Maticek et al., 1990). It consists of 71 questions (one of these has been left out of the English version of this German inventory to keep numbers in sub-scales even), designed to measure 6 variables: (1) Type 1, or the cancer-prone personality: (2) Type 2, or the CHD-prone personality: (3) Type 3, a hysterical type of personality: (4) Type 4, an autonomous, healthy type of personality: (5) Type 5, described as rational-antiemotional: (6) Type 6, an antisocial, psychopathic type of personality. The psychometric properties of these scales have been found satisfactory, (e.g. Schmitz, 1992: van der Ploeg, Kleijn, Mooks, Van Hunge Pieters & Leer, 1989), and factor analysis suggests that it might be meaningful to use a formula $(1 + 2 + 5)-(3 + 4 + 6)$ to identify probands likely to succumb to illness generally (Grossarth-Maticek et al., 1993). Scores on scales 1, 2 and 5 correlate and have consistently been associated with illness: scores on scales 3, 4 and 6 correlate negatively with 1, 2 and 5, and predict lack of illness.

Four groups form the major part of the experiment. Each consisted of 1150 persons aged from 40 to 68: subjects were consigned randomly to the group, with the sexes equally distributed. Subjects were originally chosen from electoral lists: refusal rates were around 23%. We started with altogether 4600 *S*s, of whom 3563 were tested in 1973, and followed up at the end of 1988, i.e. a I5-yr follow-up study, with death certificates as the criteria of mortality and cause of death. The groups are differentiated in terms of the method of questionnaire administration followed.

To begin with, in Group 1 *S*s had explained to them the purpose of the investigation, i.e., that the questions were concerned with an understanding of psychosocial factors in relation to health and illness. Subjects then gave informed consent to being tested, and were promised to be given results if they so wished. They were then asked to talk for 30 min about their lives, with special reference to particularly agreeable or disagreeable events. They were encouraged to talk about their typical reactions and behaviour in stressful situations. Next, the interviewer read the questions out and asked for answers. A standard explanation was associated with each question. The interviewer asked after each answer whether the question had been understood, and only if it had been answered properly was the answer accepted. If the interviewer felt that the question had not been answered in line with the explained meaning of the question, he/she repeated the question and the explanation, asking for a reconsideration. Usually it was possible to obtain a clear-cut and meaningful answer to the question after such intervention. Thus this method was designed to produce trust and understanding.

3

In Group 2 the 30-min talk at the beginning was left out, and we only have the explanatory interaction. In Group 3 there was no explanation of the questions, only the introduction concerning the purpose of the study and the nature of the 6 types. There was no discussion about positive and negative life events, but an effort was made to produce a positive reaction between interviewer and subject, and to arouse a general interest in the investigation. Finally for Group 4 there was only the presentation of the questionnaire, with the explanation of why we were carrying out the study. It was explained that no further information could be given, to preserve the objectivity of the study.

CASE HISTORIES AND SHORTER COMMUNICATIONS

Table 1. Number of Ss in the 4 administration groups, sub-divided by type. Also given are the number and proportion of individuals in each category who died of cancer and CHD

| Type | Group 1 | | | | Group 2 | | | | Group 3 | | | | Group 4 | | |
	N	CA	CHD	N	CA	CHD	N	CA	CHD	N	CA	CHD	N	CA	CHD
1	276	60 25.0%	12 4.3%	258	38 14.7%	12 4.7%	199	23 11.6%	10 5.0%	53	4 7.4%	2 3.8%
2	235	14 6.0%	28 11.9%	243	26 10.7%	30 12.3%	210	20 8.7%	18 7.8%	58	5 8.6%	3 5.2%
3	59	1 1.7%	1 1.6%	81	9 11.1%	6 7.4%	81	7 8.6%	5 6.1%	32	3 9.4%	2 6.2%
4	216	3 1.4%	2 0.9%	270	8 3.0%	3 1.1%	250	28 11.2%	16 5.2%	293	31 10.6%	14 4.8%
5	158	18 11.3%	7 4.4%	110	15 13.6%	5 4.5%	187	18 9.6%	8 5.7%	195	20 10.1%	9 4.6%
6	38	1 2.6%	1 2.6%	9	1 11.1%	1 11.1%	29	2 6.9%	2 6.9%	3	1 33.3%	1 3.3%
Total	982	106 10.8%	51 5.2%	971	97 10.0%	57 5.9%	956	98 10.1%	59 6.0%	639	63 9.9%	31 4.9%

Group 1 incorporates the usual procedures adopted in the various studies reported by Grossarth-Maticek; it embodies the elements of interviewer-subject relations we consider important in eliciting trust responses. Type 4 incorporates the usual procedures adopted in most questionnaire-type investigations. The difference is very marked, and on simple psychological grounds alone, it seems likely that Group 1 will give better results than Group 4, with Group 2 and 3 intermediate.

Results

Table 1 shows the numbers of *S*s in the four groups, and belonging to each of the 6 types. Also shown are the numbers and proportions of individuals in each group who had died of cancer or CHD (infarct). The number of participants in Group 4 is the lowest because a higher proportion of refusals occurred in that group, not perhaps unexpectedly. The other 3 groups had a much lower proportion of refusals.

Table 2 gives the percentage values for those who died of cancer in the four groups, and of the 6 types. X^2 values for 3 *df* for each Type score show that for Type 1 and Type 4 Group 1 clearly produces the best prediction, with Type 1 having the highest, Type 4 the lowest cancer mortality ($P < 0.001$). Type 6, predicted to have low cancer mortality, also falls in line with $P < 0.01$. For the other Types results are non-significant. (Here as elsewhere Yates' correction has been used when applicable.)

4

Table 3 gives results for infarct mortality. Results for Type 2 ($P < 0.05$), Type 4 ($P < 0.001$) and Type 6 ($P < 0.01$) are significant, in the expected direction. For Cancer and CHD combined, not shown in a separate table, Type 1 was significant ($P < 0.01$), Type 4 was significant ($P < 0.001$), and Type 6 was significant ($P < 0.001$).

These analyses were complicated since the number of Ss differed across each Group and Type. Analyses were conducted on the raw scores adjusted to standard raw scores in each instance so that direct, legitimate comparisons could be made. (It is of course not legitimate to undertake X^2 analyses on percentage scores, but these are given to facilitate appreciation of the trend of the data.)

We now turn to an analysis of the summed Type scores, Ill ($1 + 2 + 5$) and Well ($3 + 4 + 6$). Table 4 gives results for cancer mortality. Groups 1 and 2 produce significance levels of $P < 0.0001$ and $P < 0.001$, but groups 3 and 4 fail to show significance. Table 5 gives the values for CHD, with similar results--$P < 0.001$ and $P < 0.01$, Group 3 and 4 NS. The values for cancer and CHD combined, not shown separated, are similar.

Overall, Type 4 (the healthy, autonomous type) seems to give the best prediction of mortality-survival. X^2 values were calculated for cancer, contrasting Type 4 with Types $1 + 2 + 3 + 5 + 6$, and for CHD, also contrasting Type 4 with the rest. In both cases P values of < 0.001 were found for Group 1 and Group 2, but nothing significant for Group 3 and 4. Combining Cancer and CHD gave values of $P < 0.0001$ and $P < 0.001$ for Groups 1 and 2, NS for Group 3 and 4.

Table 2. Proportion of Ss in different groups dying of cancer

Type	Group 1. Explanation and trust	Group 2. Explanation (%)	Group 3. Trust (%)	Group 4. Questionnaires only (%)	χ^2 3 df	Significance level
1	25.0	14.7	11.6	7.4	31.36	$P < 0.001$
2	6.0	10.7	8.7	8.6	1.14	NS
3	1.7	11.1	8.6	9.4	3.11	NS
4	1.4	3.0	11.2	10.6	25.57	$P < 0.001$
5	11.3	13.6	9.6	10.7	1.19	NS
6	2.6	11.1	6.9	33.3	15.80	$P < 0.01$

Table 3. Proportion of Ss in different groups dying of coronary heart disease

Type	Group 1. Explanation and trust	Group 2. Explanation (%)	Group 3. Trust (%)	Group 4. Questionnaires only (%)	χ^2 3 df	Significance level
1	4.3	4.7	5.0	3.8	0.52	NS
2	11.9	12.3	7.8	5.2	8.83	$P < 0.05$
3	1.6	7.4	6.1	6.2	1.42	NS
4	0.9	1.1	5.2	4.9	19.71	$P < 0.001$
5	4.4	4.5	5.7	4.6	0.02	NS
6	2.6	11.1	6.9	33.3	14.54	$P < 0.01$

Table 4. Proportion of cancer deaths for Ill (1 + 2 + 5) and Well (3 + 4 + 6) types

	Ill (1 + 2 + 5) (%)	Well (3 + 4 + 6) (%)	χ^2 (1 *df*)	Significance level
Group 1. Explanation and Trust	21.8	1.6	82.38	$P < 0.0001$
Group 2. Explanation	14.1	6.0	16.30	$P < 0.001$
Group 3. Trust	10.5	9.5	0.28	NS
Group 4. Questionnaires only	7.8	11.4	1.69	NS

Table 5. Proportion of CHD deaths for Ill (1 + 2 + 5) and Well (3 + 4 + 6) types

	Ill (1 + 2 + 5) (%)	Well (3 + 4 + 6) (%)	χ^2 (1 *df*)	Significance level
Group 1. Explanation and Trust	10.2	0.9	34.62	$P < 0.001$
Group 2. Explanation	8.4	3.8	7.86	$P < 0.01$
Group 3. Trust	6.8	5.6	0.40	NS
Group 4. Questionnaires only	3.5	5.8	1.46	NS

We may now turn to look at the validity of our criterion for evaluating the procedures used by the 4 Groups. Hypothesis 1 (H_1) states that Type l predicts cancer mortality better than all other personality types. Out of 786 Type 1 *S*s, 134 died of cancer, i.e. 17%. Of the other types, 231 died of cancer, out of 2777, i.e. 8.3%. X^2 with one *df* = 23.62, $P < 0.001$. H_1 is supported.

H_2 states that Type 2 predicts CHD mortality significantly more often than all the other personality types. Out of 766 Type 2 *S*s, 79 died from CHD, i.e. 10.3%. Out of 2797 other Types, 119 died from CHD, i.e. 4.3%, giving a X^2 value of 19.29, $P < 0.001$. H_2 is supported. H_3 states that Type 1 predicts cancer death significantly more often than does Type 2. Of Type 1 *S*s, 134 out of 786 *S*s died of cancer, i.e. 17%. Out of 766 Type 2 *S*s, only 65 died of cancer, i.e. 8.5%, $X^2 = 22.00$, $P < 0.001$. H_3 is supported. H_4 states that Type 2 predicts CHD better than Type 1. Out of 766 Type 2 *S*s, 79 died of CHD, i.e. 10.3%. But out of 786 Type l *S*s, only 36 died of CHD, i.e. 4.6%. $X^2 = 16.89$, $P < 0.001$. H_4 is supported. And finally, H_5 states that Types (3 + 4 + 6) are significantly less likely to die from cancer or CHD than Types (1 + 2 + 5). Of 1361 Types (3 + 4 + 6), 149 died of cancer or CHD, i.e. 10.9%. But of Types (1 + 2 + 5), 414 died from cancer or CHD, out of 2202 *S*s, i.e. 18.8%. $X^2 = 28.21$; $P < 0.001$. H_5, too, is supported.

6

Discussion

It will be clear that our results on the whole bear out the general theory they were meant to test. The research was premised on the hypotheses numbered 1 to 5, namely that (H₁) Type 1 would predict cancer, (H₂) Type 2 would predict CHD, (H₃) Type 1 would predict cancer better than Type 2, (H₄) Type 2 would predict CHD better than Type 1, and (H₅) that Types 3 + 4 + 6 would die less frequently from cancer or CHD than Types 1 + 2 + 5. All 5 hypotheses were supported at the $p < 0.001$ level, and are hence suitable to form the criterion against which to test our hypothesis that *method of data collection would significantly affect validity of findings*.

It is clear that Group I (explanation plus trust) does better than the other Groups, while Group 4 (neither) does worst. This is in line with prediction. Groups 2 and 3 are between Groups 1 and 4; this too is in line with prediction. No prediction was made as to the relative performance of Groups 2 and 3. It is evident that Group 2 (explanation) does better than Group 3 (trust); indeed, Group 3 performs only marginally better than Group 4. As an example, consider the differences between Ill (1 + 2 + 5) and Well (3 + 4 + 6) *S*s, derived from Table 4. Details are given in Table 6. Clearly, explanation was more important than trust in producing veridical predictions, although of course the combination seems to have synergistic properties. If we simply add the effects of explanation (Group 2) and trust (Group 3), we get figures below those achieved by Group 1, where explanation and trust were combined in the method of administration. Thus for cancer the figures are 20.3 vs 9.8; for CHD, 9.3 vs 5.6, and for both combined 29.7 vs 15.5. Thus trust and explanation condition seem to interact synergistically, i.e. their effects multiply rather than add. This is an important finding. Of course these figures cannot necessarily be extrapolated to other questionnaires, which might have less complex wording. Much research remains to be done along these lines to discover the general laws linking properties of the questionnaire and properties of administration procedure.

Table 6. Differences in mortality between Well and Ill groups for different questionnaire administrations

	Cancer	CHD	Both
Group 1. Explanation and trust	20.3	9.3	29.7
Group 2. Explanation	8.8	4.6	13.4
Group 3. Trust	1.0	1.0	2.1
Group 4. neither	−3.6	−2.3	−5.9

Table 7. Proportion of Ss dying of cancer and CHD, when diagnosed as types 1 to 6

	N	Group 5 Cancer	CHD
Type 1	395	68 17.2%	12 3.0%
Type 2	347	10 2.9%	21 6.1%
Type 3	152	3 2.0%	3 2.0%
Type 4	335	5 1.5%	4 1.2%
Type 5	268	36 13.4%	12 4.5%
Type 6	72	2 2.8%	2 2.8%
Total	1569	124 7.9%	54 3.4%

It will have been noted that for Group 4 results are actually negative, i.e. the "ill" Ss (Types 1 + 2 + 5) actually have less cancer and CHD than the "well" Ss (Types 3 + 4 + 6). This is in line with the theory that it is suppression of emotion that is mainly responsible for cancer and CHD. This hypothesis was first tested by Kissen and Eysenck (1962), who found that as predicted cancer patients had significantly lower scores on a neuroticism questionnaire than patients with benign growths. Many other studies have since replicated this finding (Eysenck, 1985). The negative relationship between the Types 1 + 2 + 5 combination and illness is well in line with this view. Presumably the procedures used in Groups 1, 2 and 3 overcome to varying degree the suppressive proclivities of cancer-prone and CHD-prone Ss, and succeed in turning negative into positive values in Table 6. This is another important area of research. We need to classify the conditions under which disease-prone Ss suppress the verbal expression of their emotions, as opposed to verbalizing them (Eysenck, 1991). Many negative findings may be due to a failure to use appropriate methods of investigation; clearly contrary results may seem to appear from different modes of data collection.

EXPERIMENT 2

Few experimenters in a field like epidemiology that requires large follow-up groups can afford the expensive use of interviewer-administered testing procedures that produced such good results in Experiment 1. Experiment 2 was conducted to investigate the possibility of obtaining reasonably good results with a procedure that would be less expensive in interviewer time. In this experiment, (1) we used a new set of questionnaires written specially to be easier to understand, (2) we dispensed with the trust-creating interaction part, and (3) minimized the explanation aspect by only offering explanation when asked specifically. This study also used a random sample, half male, half female; 1800 persons, aged between 40 and 67, were approached by letter, and 1569 agreed to fill in the questionnaires. The study began in 1976-1978, so that we have a 10-year follow-up, i.e. 5 yr shorter than the follow-up in Experiment 1.

Four questionnaires were used, the first of which gave the Ss scores for the 6 Grossarth-Maticek types, as in Experiment 1. The second asked questions about self-regulation, an extended list of questions using the concept of autonomous behaviour basic for Type 4. The third questionnaire

related to contentment, while the fourth contained additional questions to differentiate Type 1 from Type 2. None of these have been published, but publication is due in the near future.

The method of administration was as follows. In the introduction, the interviewer explained the aim of the study. Next, the questionnaire was read out to the *S* who was asked to explain his reasons for giving a specific answer. Only when the question was obviously misunderstood would the interviewer repeat the question and ask for a new explanation of the *S*'s answer. The results of the study are shown in Table 7.

Clearly Type 1 is linked with cancer, Type 2 with CHD, with the former link again much stronger. Type 5, too, is linked with cancer, less so with CHD, while Types 4 and 6 show no difference. Statistics have been worked out as before for the "ill" and "'well" combination of Types 1 + 2 + 5 vs Types 3 + 4 + 6. For cancer the percentages are 14.1 vs 1.7%, $P < 0.001$, and for CHD they are 6.1 vs. 1.4%, $P < 0.001$. Again results are significant in the predicted direction, and again the figures for cancer are much better than for CHD. X^2 for cancer is 83.06, for CHD it is 25.03, with 1 *df.*

The figures for Type 4, contrasting this with the sum of the other 5 Types, is 1.5 vs 9.6% for cancer, and 1.2 vs. 4.1% for CHD, both with $P < 0.001$. Clearly, it is possible to obtain good results even with a somewhat curtailed methodology, although of course even this curtailed methodology is more demanding than the usual habit of just handing out questionnaires.

GENERAL DISCUSSION

Many commentators have remarked on the contradictory results often reported in the literature concerned with the effects of psychosocial factors (stress, personality) on cancer and CHD. This and the preceding study (Grossarth-Maticek et al., 1993), as well as the experience of Type A-Type B investigators (Eysenck, 1990a) of the superiority of using a structured interview rather than simply handing out questionnaires, suggest that one important variable leading to disparate results may be the method of data collection. Motivation, general attitude, awareness of the purpose of the study, understanding of the questions, and personal involvement--these may be crucial in determining the answers a person will give to questions that touch upon very sensitive, often painful, always very personal aspects of his behaviour. It is naive to disregard these factors and imagine that truthful answers will always be given even when no measures are taken to ensure co-operation, trust and understanding of the questions. The practice of using students as subjects may have had a strong influence here; students are more likely than ordinary people to appreciate the purpose of a scientific study, to agree with the need for such studies, and to understand the questions properly. Nothing of the kind can be presumed in non-student populations, and the very frequent misunderstandings of questions reported by our interviewers is eloquent testimony to such needs. The Schmale and Iker (1971) study is eloquent witness to the need for appropriate methodology and instruments in this field; they succeeded in verifying their theory that hopelessness was a valid prediction of cancer when they used an appropriate structured interview, but failed dismally (using the same subjects) when they used the MMPI and the Rorschach (Eysenck, 1990b). Measurement in epidemiology needs to be informed on psychological issues if results are to be taken seriously (Eysenck, 1991).

9

REFERENCES

Amelang, M. & Schmidt-Rathjens, C. (1992). Personality, stress and disease: Some results on the psychometric properties of the Grossarth-Maticek and Eysenck inventories. *Psychological Reports*, 71, 251-263.

Campbell, D, T. (1960). Recommendations for APA test standards regarding construct, trait or discriminant validity. *American Psychologist*, 15, 546-553.

Cronbach, L., & Meehl, P. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-302.

Eysenck, H. J. (1985). Personality, cancer and cardiovascular disease: A causal analysis. *Personality and Individual Differences*, 6, 535-556.

Eysenck, H. J. (1990a). Type A behaviour and coronary heart disease: The third stage. *Journal of Social Behaviour and Personality*, 5, 25-74.

Eysenck, H. J. (1990b). The prediction of death from cancer by means of personality/stress questionnaire: Too good to be true? *Perceptual and Motor Skills*, 71, 216-218.

Eysenck, H. J. (1991). *Smoking, personality and stress: Psychosocial factors in the prevention o/" cancer and coronary heart disease*. New York: Springer.

Eysenck, H. J. & Fulker, D. (1983). The components of type A behaviour and its genetic determinants. *Personality and Individual Differences*, 4, 499-505.

Garber, J. & Strassberg, Z. (1991). Construct validity: History and application to developmental psychopathology. In Grawe, W. & Ciachetti, D. (Eds), *Personality and psychopathology* (pp. 219-258). Minneapolis: University of Minnesota Press.

Grossarth-Maticek, R. & Eysenck, H. J. (1990). Personality, stress and disease: Description and validation of a new inventory. *Psychological Reports*, 66, 355-373.

Grossarth-Maticek, R., Eysenck, H. J. & Barrett, P. (1993). The prediction of cancer and coronary heart disease as a function of the method of questionnaire administration. *Psychological Reports*, 73, 943-959.

Jenkins, C. D., Zyzanski, S. J., Ryan, T. J., Flessas, A. & Tannenbaum, S. (1977) Social insecurity and coronary-prone type A responses as identifiers of severe atherosclerosis. *Journal of Consulting and Clinical Psychology*, 45, 1060-1067.

Kissen, D. M. & Eysenck, H. J. (1962). Personality in male lung cancer patients. *Journal of Psychosomatic Research*, 6, 123-137.

Lee, P. N. (1988). *Misclassification of smoking habits and passive smoking: A review of the evidence*. Berlin: Springer.

Loewinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports*, 3, 635~599.

van der Ploeg, H., Kleijn, W. C., Mooks, J., van Hunge, M., Pieters, A. & Leer, J.-W. (1989). Rationality and anti-emotionality as a risk factor for cancer: Concept differentiation. *Journal of Psychosomatic Research*, 33, 212-225.

Schmale, H. A. & Iker, H. (1971). Hopelessness as a mediator of cervical cancer. *Social Science and Medicine*, 5, 55-100.

Schmitz, P. (19923. Personality, stress reaction and disease. *Personality and Individual Differences*, 13, 683-691.

Temoskok, L. & Dreher, H. (1992). *The Type C connection*. New York: Random House.