(2)

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
31 October 1991

3. REPORT TYPE AND DATES COVERED
Annual, 1 Sept, 1990 - 31 Aug, 1991

4. TITLE AND SUBTITLE
The central executive component of working memory

5. FUNDING NUMBERS
G. AFOSR-90-0343
PE- 61102F
PR- 2313
TA- A4

6. AUTHOR(S)
A. Baddeley., J. Duncan and H. Emslie

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
MRC Applied Psychology Unit
15 Chauncer Road
Cambridge CB2 2EF

8. PERFORMING ORGANIZATION REPORT NUMBER
AFOSR-[illegible]- 91 - 1006

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
EOARD
223/231 Old Marylebone Road,
London NW1 5TA, U.K.

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This research is based upon the hypothesis that three different phenomena - behavioural impairments after frontal lobe damage, "general intelligence" or Spearman's g, and interference between dissimilar concurrent tasks -all reflect the operation of a central executive (CE) system involved in the organisation of many different kinds of behaviour. Four sets of experiments are presented. One set shows that frontal lobe damage produces massive impairments in "intelligence tests" based on current problem-solving ability. A second shows that one characteristic frontal error - mismatch between knowledge of a task's requirement and the resultant behaviour - can also be reliably produced in normals, and is closely related to g. The third set of experiments is based on the idea that executive processes lose importance as behaviour becomes stereotyped or automatic. If so, generating random sequences should load the CE, whatever their particular content, and the experiments indeed suggest that the demands of random generation are similar for verbal and manual materials. Similarly, the fourth set of experiments suggests that correlations between reaction time and g diminish with practice only if there are no switches in mental set. It is proposed that the CE is a system for detection/selection of goal states in novel behavioural settings.

14. SUBJECT TERMS
Working memory, central executive, frontal lobes, intelligence

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	(U)

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to ***stay within the lines*** to meet ***optical scanning requirements***.

Block 1. Agency Use Only (*Leave blank*).

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C -	Contract	PR -	Project
G -	Grant	TA -	Task
PE -	Program Element	WU -	Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. *(If known)*

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

- DOD - See DoDD 5230.24, "Distribution Statements on Technical Documents."
- DOE - See authorities.
- NASA - See Handbook NHB 2200.2.
- NTIS - Leave blank.

Block 12b. Distribution Code.

- DOD - Leave blank.
- DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
- NASA - Leave blank.
- NTIS - Leave blank.

Block 13. Abstract. Include a brief *(Maximum 200 words)* factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code *(NTIS only)*.

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

Introduction

Since the original work of Baddeley and Hitch (1974), converging methods have been used to distinguish separate components of the working memory system. These include the articulatory loop - storing 1-2 seconds' of inner speech - the visuospatial scratchpad, and various specific motor buffers. Perhaps strangely, less attention has been given to the most central component of Baddeley and Hitch's (1974) scheme - the central executive (CE). In this project we are again using converging methods to analyze CE function.

Three phenomena form the basis for our work. First is the effect of damage to the frontal lobes of the brain. Characteristically such damage results in a widespread disorganisation of behaviour; instead of a goal-directed sequence, behaviour can appear to be "fragmented", "bizarre" or "irrelevant" (Luria, 1966). Importantly, such disorganisation influences behaviour in many different realms, from perception and memory to regard for social conventions. By hypothesis, this widespread effect reflects a CE impairment.

Second we are interested in individual differences in the normal population, in particular evidence for a factor of "general intelligence" or Spearman's g. Across individuals, even tasks that appear superficially dissimilar will usually show some positive correlation - to some extent, people who do better on one of the tasks will tend also to do better on the other. According to the g hypothesis, such widespread positive correlations indicate the existence of a general or g factor, making some contribution to all manner of different tests. By hypothesis, again, this general system is the CE. Accepting the g factor interpretation, it is easy with factor analysis to show which tests are most strongly correlated with g, i.e. which are the best measures of an individual's CE function. This is the basis for design of standard "intelligence" or IQ tests such as the WAIS, Raven's Matrices etc.

The third consideration is dual task interference. When two tasks are carried out together, interference between them depends partly on their similarity in e.g. input modality, output madality, or information content. For example, there will often be strong interference between two verbal tasks, suggesting shared demands on the articulatory loop. Even dissimilar tasks, however, usually show some interference. Again by hypothesis this may reflect shared demands on the CE.

Our project has two related goals. First, we ask whether the phenomena of frontal lobe damage, Spearman's g, and interference between (dissimilar) concurrent tasks really can be linked empirically. Second, if indeed we obtain converging evidence for a CE system, we should like to ask what is its exact processing function, i.e. what contribution it makes to the organization of behaviour.

It may be useful to begin with a summary of our findings to date, and the working hypothesis that they suggest. Various lines of evidence suggest that indeed frontal lobe deficit is very closely related to Spearman's g. Factors that influence a task's g involvement include the number of separate requirements or goals, the degree of consistent practice, and the nature of environmental cues or prompts to the required behaviour. It is a commonplace that human behaviour is intrinsically goal-directed, raising the question of how goals are selected, among many competing alternatives, for control of current behaviour. Our provisional suggestion is that the CE has a central role in goal detection/selection in unfamiliar behavioural settings. There are four sets of experiments to be considered, dealing respectively with (i) generation of random sequences, (ii) performance of frontal patients on standard intelligence tests, (iii) goal selection, and mismatch between knowledge of a task's requirements and the corresponding behaviour, (iv) consistency of mental set and the relationship of g to choice reaction time.

Random generation

It has often been suggested that executive functions are most important in the early phase of skill development, before practice has rendered behaviour "automatic" (e.g. Bryan & Harter, 1899; Norman & Shallice, 1980). Correspondingly, practice has been reported to reduce frontal lobe deficits (Luria & Tsvetkova, 1964), g correlations (Ackerman, 1988) and dual task interference (Schneider & Shiffrin, 1977). With this in mind, Baddeley (1986) has suggested that an excellent task for loading the CE may be generation of random sequences. In this task, the basic requirement is exactly to avoid consistent, familiar or stereotyped behaviour, i.e. the sort of behaviour that may easily be automatized. Because of the minimal stimulus input, furthermore, random generation may be an easy task to use in many dual task settings. Indeed, prior work has shown its interfering effects on concurrent tasks ranging from chess (Baddeley, in press) to driving (Duncan, Williams, Nimmo-Smith, & Brown, in press).

This prior work, however, is far from establishing random generation as a task loading any general-purpose CE system. Typically the task has been to generate sequences of spoken letters or digits in time with a metronome. In our first two experiments, we wished to examine the relationship between this verbal task and a manual equivalent. In part the reason is practical - in some dual task experiments, it may be important to distinguish interference due to shared CE requirements from that due to a shared verbal component. The experiments also address theoretical issues, however. A characteristic of letter and digit responses is that there is a strong population stereotype (alphabet, counting) to be resisted - is this an important component of the random generation task? If so, manual tasks with less obvious population stereotypes might give very different results. In the task as it has usually been used, candidate responses have to be generated entirely from memory. To assess the impact of this factor, we compared conditions with and without explicit visual cues to the response set.

In Experiment 1 we compared three different random generation tasks. In the first, responses were the spoken digits 0 to 9. In the second, they were keypresses made with the ten digits of the two hands. A special keyboard, interfaced to a Macintosh computer, was built so that each digit rested comfortably on a separate key, with barriers to prevent slippage between one key and the next. In the third task only four of these keys were used, operated with the first and second fingers of the two hands. This time each response was a chord; legitimate chords were all possible two- and three-finger combinations, again giving ten response alternatives in total. In each task, responses from the ten-alternative set were to be generated in random order, in time with a metronome beating at one of 3 different rates. For digit and single-key tasks we used rates of 0.5, 1.0 and 1.5 sec/response; for the harder chord task we used 1.0, 1.5 and 2.0 sec/response. Thus the experiment involved a total of 9 conditions; each subject produced 2 blocks of approximately 120 responses per condition, in a single session of about 1 1/2 hours, with the order of conditions counterbalanced across the whole group of 18 subjects.

In Experiment 1 there were no visual cues to the response set. In the manual tasks, in particular, the hands and keyboard were hidden using a framework built of cardboard. Experiment 2 involved the same tasks but with visual cues intended to help generation of candidate responses. In the digit task the digits 0-9, written out in a row, were available to the subject's view. In manual tasks the keyboard was now visible.

Success in producing a random sequence was scored in several different ways; as in previous research (Baddeley, 1966), these gave rather similar results. Accordingly in Figure 1 we present only one measure - per cent redundancy in digram frequencies. Derived from information theory, this measure varies from an ideal of zero (obtained if each possible response digram is used equally often) to a theoretical maximum of one hundred (only a single response digram ever used). The higher is the score, the stronger is the bias towards certain favourite response digrams, and thus the greater the deviation from randomness.

Insert Figure 1 about here

Three findings may be noted. First, all three tasks showed an increase in randomness with increased time per response. Thus this characteristic finding for verbal tasks (Baddeley, 1966) is replicated and perhaps even strengthened in the manual tasks. Second, in terms of attained randomness, the easiest task was digits and the hardest chords, with single keys intermediate. There is no suggestion here that resisting a strong population stereotype makes the digit task especially difficult. Third, visual cues to the response set made all tasks a little easier, but left the pattern of data otherwise unchanged. The key factor appears to be a requirement for random behaviour per se, rather than any difficulty in retrieving or generating candidate responses.

Though these results are certainly consistent with a general-purpose CE, involved in both verbal and manual random generation, a more direct test is planned in our next experiment. In this study verbal and manual tasks will be carried out concurrently. Very substantial interference is predicted between the two.

The last findings from these experiments concern individual differences. If different random generation tasks all load a common CE system, do they also provide a good measure of individual differences in CE efficiency? If so they should show strong correlations both with one another and with more standard "intelligence" tests. Here our results were negative. Across experiments, the mean correlation between scores in verbal and manual tasks was only .39, while between single-key and chord tasks it was .50. In a further study of 41 elderly adults (see below), we obtained only negligible correlations between random digit generation and a standard test of g (Culture Fair; Institute for Personality and Ability Testing, 1959).

To sum up, we suspect that random generation tasks using verbal and manual responses indeed load a common CE system. In support of this, a recent PET scanning study has shown activation of area 46 in dorsolateral frontal cortex in both verbal and manual generation tasks (Frith, Friston, Liddle, & Frackowiak, 1991). As a measure of individual differences in CE efficiency, however, random generation seems not to be a promising candidate.

<u>Performance of frontal patients in IQ tests</u>

While others have proposed links between frontal dysfunction and dual task interference (Norman & Shallice, 1980), and between dual task interference and g

(Ackerman, 1988), it is conventionally stated that frontal dysfunction has little to do with the kind of "intelligence" measured by standard tests. This belief derives from the finding that frontal patients often show little evidence of impairment in IQ (e.g. Hebb & Penfield, 1940; Mettler, 1949; Weinstein & Teuber, 1957).

In interpreting such results there are two things to bear in mind. First is a distinction between different kinds of frontal lesion. Quite probably, the kind of widespread disorganisation of behaviour that we are attributing to CE impairment simply does not occur in many cases of circumscribed lesion affecting only a small part of the frontal lobe. Such patients may show deficits in few if any aspects of behaviour. In our work, however, we have focussed on a second point. This is a distinction between two different types of IQ test, based on entirely different principles of measurement.

The first type of test works by measuring a person's average score on a diverse set of sub-tests, each of which on its own may have only a small correlation with g. As Spearman (1927) showed, the method works with any reasonably diverse set of sub-tests, for much the same reason as averaging 100 poor estimates of any quantity produces a single much better estimate. Though this is the basis of the most clinically popular tests like the WAIS, it has obvious theoretical disadvantages. In particular it is based on individual measures of behaviour to which any "g" system is making only a small contribution.

The second approach is to find single tests which on their own have a substantial g correlation. Commonly these turn out to be tests with a substantial element of novel problem-solving, e.g. matrices, verbal analogies. Here then we have a kind of performance which is heavily dependent on any "g" system (by hypothesis, the CE).

In Experiment 3, carried out in collaboration with Paul Burgess at the University of London, we tested 3 patients of the kind specifically taken to show that frontal impairment is unrelated to conventional "intelligence". All had a major frontal lesion and substantial clinical evidence of a frontal behavioural impairment;

on the WAIS, however, their IQs ranged from 126 to 130. Our interest was in IQs measured by the second, problem-solving method. The test we used was Cattell's Culture Fair (Institute for Personality Testing, 1959), which involves problem-solving of various kinds with simple visuospatial materials.

Results are shown in Table 1. The average discrepancy between WAIS and Culture Fair IQ was 30 points. Thus we have strong evidence that frontal impairments are associated with substantial losses in performance in a test which is very heavily dependent on any "g" system.

Insert Table 1 about here

Data from 7 other patients - tested in collaboration with Roger Johnson at Addenbrooke's Hospital - support this conclusion. This time we did not select patients for preserved WAIS scores; these were simply head injured patients with CT evidence for damage largely confined to the frontal lobes. For each patient, two or three control subjects were chosen, matched for age, sex, socioeconomic group, and performance on the National Adult Reading Test (Nelson, 1982). As measured on the Cattell Culture Fair, mean IQ in the patient group was 11 points lower than in the controls. Again, this test - a problem-solving test which we should expect to make heavy demands on any "g" system - reveals clear losses associated with frontal lobe damage.

Mismatch between knowledge and behaviour

One of the most striking results of frontal impairment is the occasional observation of mismatches between knowledge of a task's requirements and the resulting behaviour. Though the patient shows verbally that instructions have been understood, behaviour suggests no attempt to satisfy them. As an example, Luria (1966) describes patients asked to raise their hand in response to a light; when the light came on, the patient might say, "I must raise my hand!" yet make no attempt to do so. Similarly, Baddeley (1986) described a patient who was given scissors and string but asked not yet to cut. He began at once to do so, and when the error was pointed out, continued while at the same time saying, "Yes, I know I'm not to cut it."

Our interpretation of this phenomenon concerns the problem of goal selection. At any given time, a person must select which goal or goals to pursue, among many potential alternatives. This issue has been addressed mainly in the literature on problem-solving; in this context, several people have pointed out that we need some scheme like the one shown in Figure 2 (e.g. Anderson, 1983; Duncan, 1990; Duncker, 1945). One set of candidate goals is suggested by current events in the environment (current state), suggesting potential next-states that it might be useful to pursue. This process is needed so that new environmental input can always overturn current concerns; for example, if one enters the kitchen intending to cook breakfast but finds that the cat has brought in and released a bird. Another set of candidates is suggested by currently active goals; in Figure 2 these are called relevant next-states, i.e. states that would bring an active goal closer ("sub-goals"). This entire set of candidate next-states must be weighted by net "importance", the most important then being chosen for pursuit.

Insert Figure 2 about here

In this context, verbal instructions may be seen as one form of environmental input specifying candidate goals ("task requirements"). At least in the context of an experiment, furthermore, there is a strong expectation that such goals will be so highly weighted as to take control of behaviour. When this does not happen in the frontal patient, one has direct evidence for an impairment in the goal selection process.

A problem for experimental work on this phenomenon is that in the past it has been reported only as an occasional, unpredictable event. What we need is a task in which the phenomenon occurs more frequently. In our grant proposal we described a promising candidate. Pairs of letters or numbers are presented side by side in the centre of an APPLE II screen. On each trial thirteen pairs are presented, one after the other, at a rate of 400 msec/pair (200 msec presentation, 200 msec interstimulus interval or ISI). There are three basic requirements. First, the task is to repeat aloud any letters, ignoring numbers. Second, the subject watches for letters on only one side at a time, left or right. An instruction WATCH LEFT or WATCH RIGHT appears at the start of the trial, before the stimulus sequence begins. Only letters from the specified side are to be repeated. Third, there is a cue between the tenth and eleventh pairs which may require a switch of sides. This cue is a symbol, either + or -; a + means that for the remainder of the trial the subject should watch right, while a - means watch left. The cue is presented in the centre of the screen, with the same timing as the stimulus pairs (200 msec presentation, separated from preceding and following pairs by 200 msec ISIs) to minimise disturbance in the visual sequence. Altogether there are always 7 letter pairs per trial, two of which occur after the +/- cue, in positions 12 and 13. Stimuli from a sample trial are shown in Figure 3.

Insert Figure 3 about here

In pilot work we found that the first two task requirements - to repeat letters while ignoring numbers, and to begin on the correct side - were always faithfully reflected in performance. Even a few normal subjects, however, showed no evidence of an attempt to satisfy the third requirement, response to the +/- cue - even though all could successfully describe this requirement. Here we present data from 4 experiments: (a) a basic group of 90 adults, tested before the beginning of the present grant; (b) a group of 7 frontal patients; (c) a group of 41 elderly people; (d) a group of 38 adults doing the basic task concurrently with a dot-key reaction time (RT) task.

Experiment 4

Subjects in this experiment were 90 normal adults, mean age 41, range 29-47. They were obtained through a local employment agency.

Each subject was given 3 blocks of 12 trials. Each block consisted of 3 four-trial subblocks, which we term subblocks a (trials 1-4), b (5-8) and c (9-12). Thus each subject had 9 subblocks altogether, 1a to 3c. Within each subblock, each possible combination of starting and finishing sides was used once (in random order): start left, stay on left (- cue); start left, switch to right (+ cue); start right, stay on right (+ cue); start right, switch to left (- cue). A trial was counted as "passed" if, both before and after the +/- cue, more letters were reported from the correct than from the incorrect side. Even a subject who was ignoring the +/- cue would typically pass on half the trials (the "stay" trials on which no switch of side was required). Thus our main score was based on passing or failing whole subblocks; a subblock was "passed" if it had at least 1 passed "stay" trial and 1 passed "switch" trial.

Our main data come from block 1. Table 2 shows the distribution of number of passed subblocks, across the 90 subjects. It is immediately apparent that subjects fell into two groups. The majority (73/90) picked the task up fairly readily, passing either 2 or 3 subblocks. There were 15 subjects, however, who failed all 3 subblocks, and on closer examination all 15 could be seen to be ignoring the +/- cue altogether. For 14 of the 15 the strategy was simply to stay on the side indicated at the outset of the trial.

Insert Table 2 about here

A first question is whether these people had understood and remembered the instruction. Several measures were taken to ensure that they had. After an initial practice trial (before the start of data collection in block 1), subjects were asked, "What do you do when you see a +? and a -?" If the answer was incorrect, the rule was repeated and further practice trials given, until the question was answered correctly. As an aid to memory, furthermore, two pieces of paper were left permanently on the table, one with the word MINUS placed to the subject's left, and the other with the word PLUS to the right. To check on the efficacy of these measures, each subject was asked again to repeat the rule at the end of block 1. All subjects did so correctly.

A second question is whether subjects were simply unable to obey the rule, e.g. because they had insufficient time to switch sides. Here the data from block 2 are relevant. Between blocks 1 and 2, as we have said, every subject was asked to repeat the +/- rule. For half the subjects, furthermore, additional emphasis was given to this aspect of the task by asking, after each trial in Block 2, which cue had been seen, which side had been watched, and whether or not performance had been correct. The result was that 7/15 subjects then passed subblock 2a or 2b and proceeded to complete the rest of the task correctly. Apparently the problem for these subjects was not that they were unable to obey the initial instruction, but that this instruction on its own was insufficient to ensure that the third task requirement - response to the +/- cue - would be adequately focussed upon. Thus further emphasis on this aspect of the task caused the initial failure to be corrected.

This point may be put more formally. Only very rarely did any subject fail a subblock after his or her first pass. Thus performance could be described as a series of 0...n initially failed subblocks, followed by correct performance. In Table 3 (first row) is shown the probability, for each subblock 1a to 3c, of a correct pass given no prior pass. Formally this is the hazard function for first correct pass. For example, of the 90 people beginning the task, 60 (.67) passed the first subblock (1a). Of the 30 who failed, 14 (.47) then passed the next subblock (1b) - and so on for the remainder of the task. The findings may be summarised as follows. The hazard function decreased steeply through block 1. By subblock 1c, in particular, there was almost no chance that a subject who had failed 1a and 1b would now pass spontaneously. The hazard function increased again, however, at the start of block 2, when prompts had been given to focus attention on response to the +/- cue. Though this phenomenon was stronger in those subjects who were asked questions about their performance after each block 2 trial, it occurred even in those asked simply to repeat the rule after block 1.

Insert Table 3 about here

A final finding concerned the relationship of this phenomenon to scores on the Culture Fair test, also given to all 90 subjects. For this purpose, block 1 as a whole was simply scored "pass" (at least 1 correct subblock) or "fail". In Figure 4, probability of a "fail" is shown as a function of performance on the Culture Fair, expressed as a z-score based on the published norms. A Culture Fair score of -0.5, for example, means a score lying half a standard deviation below the mean of the normal population. Also shown at the bottom of the figure are numbers of subjects falling into each bin.

Insert Figure 4 about here

The results look extremely promising. Failure on our task was almost certain for subjects whose Culture Fair scores were more than one standard deviation below the mean. It never occurred for subjects more than one standard deviation above the mean. It should be noted, however, that very few subjects fell in these extreme groups. This is why in total there were only 15/90 failures; correspondingly, though the relationship between our task and the Culture Fair seems in Figure 4 to be so strong, the product-moment correlation between pass/fail (scored 0/1) and Culture Fair errors was only .38.

To sum up: We have here an error which seems closely related to the characteristic frontal mismatch between knowledge and behaviour. Though a task requirement has been understood - and though the subject may be quite capable of satisfying this requirement - behaviour suggests no attempt to do so. Though the rarity of the phenomenon makes conclusions hard to draw, furthermore, there is a strong suggestion that its occurrence is closely related to Spearman's g.

Experiment 5

In Experiment 5 we sought direct evidence that frontal lobe damage would produce neglect of the +/- cue in our task. We tested the same 7 patients described earlier - head injury patients selected for CT evidence of damage that was confined largely to the frontal lobes. Since various different cueing procedures were tried in blocks 2 and 3, only the data from block 1 are presented.

Details of the patients and pass/fail (0/1) scores are shown in Table 4. Five of the seven patients failed, as compared with 15/90 normal adults in Experiment 1.

Insert Table 4 about here

Experiment 6

In Experiment 6 we had three goals. First, we wished to test more people with low scores on the Culture Fair test, to confirm the rather strong suggestion in Experiment 4 that such people are extremely likely to neglect the +/- cue in our task. A standard finding is that, with advancing age, performance on problem-solving tests like the Culture Fair drops off rather substantially, while performance on more knowledge-based IQ tests like vocabulary is preserved (e.g. Cattell, 1971). Accordingly we selected a group of 41 normal elderly adults, aged between 60 and 70, drawn from the paid subject pool of the Applied Psychology Unit.

Second, we wished to know whether any means of putting emphasis upon the +/- cue would cause subjects who had failed block 1 to pass in block 2. Half the subjects were given the same emphasis procedure as before - following a question about the rules at the end of block 1, they were asked after each trial of block 2 which cue had been seen, which side had been attended, and whether performance had been correct. The other half of the subjects were not asked about the rules or to comment upon their performance. Instead, on each trial of block 2, large arrows appeared without warning along with the cue, above and below and pointing in towards it. To check that the rules had actually been remembered, these subjects were asked to repeat them at the end of block 2.

Lastly we wished to know whether, given sufficient emphasis, all subjects would prove capable of correct cue use. In block 3, any errors in cue use were explicitly pointed out by the experimenter, and the subject was asked to try and correct them.

There were also some subsidiary goals. According to our hypothesis, a test like the Culture Fair - with its strong g saturation - should be the best predictor of errors in our task. Another possibility is that we are measuring some specifically verbal deficit, since, after all, the phenomenon is a failure to use information in verbal instructions. To test this idea we gave three other tests related to specifically verbal intelligence: the Mill Hill test of vocabulary (Raven, Raven & Court, 1988), the Nelson-Denny test of reading comprehension (Brown, Nelson & Denny, 1976), and the Daneman-Carpenter test of sentence span (Daneman & Carpenter, 1980).

The first result was that, as we should expect, selection of elderly subjects was successful in producing a distribution of somewhat poor Culture Fair scores along with preserved Mill Hill scores. Mean Culture Fair and Mill Hill IQs were 91 and 107 respectively. Though it is not a point we wish to pursue here, such standard findings strongly suggest that declining abilities in old age are strongly associated with deficits in the CE (Baddeley, 1986).

The second result was a clear replication of our prior findings on cue neglect and g. The data are shown in Figure 5, which this time is truncated at a Culture Fair score of +0.5 since no subject scored above this value. In this experiment, the correlation between pass/fail on block 1 and Culture Fair errors was .54. Again, even those subjects who failed always correctly described the +/- rule when asked.

Insert Figure 5 about here

The third result concerned the hazard function for first correct pass, shown in the second row of Table 3. This time, the decline through block 1 was less clear, since these elderly subjects were having more trouble picking the task up quickly. Accordingly the increase in subblock 2a was not significant, though it did occur for subjects receiving both verbal and arrow cue emphasis.

The fourth result is also shown in the hazard function. When explicit verbal error feedback was initiated in block 3, even subjects who had previously failed began now to pass, all subjects in fact succeeding by subblock 3c. This finding confirms that even the poorest subjects were actually capable of correctly using the +/- cue. Their difficulty lay rather in focussing attention upon this particular task requirement.

Lastly, the results confirmed that neglect of the +/- cue was related more closely to the Culture Fair than to our other, verbal IQ measures. Correlations with Mill Hill, Nelson-Denny and Daneman-Carpenter were .05, .39 and .23 respectively. The low correlation with Mill Hill shows especially clearly that, in elderly subjects, the key factor is the current level of g - as measured by current problem-solving ability - not verbal skills acquired earlier in life, before the changes of old age took place.

Experiment 7

Our results show that the process of using verbal instructions to establish goals for performance has much in common with the kind of "problem-solving" required in the Culture Fair. Analysis of a typical "Culture Fair" item (Figure 6) suggests why this should be. Note that to protect confidentiality this item has been made up to resemble those in the test, without actually being drawn from it. The task is to choose which of the numbered boxes on the right would fit correctly into the empty cell of the matrix on the left.

Insert Figure 6 about here

Just like a set of verbal instructions, one might say that the stimulus materials for this item provide a set of cues to goals that should be satisfied in performance (Carpenter, Just, & Shell, 1990). Thus the choice made must satisfy variations in shape, in elongation, and in shading. Each of these goals must be detected from the stimulus materials, selected and satisfied. Beyond this, there are at least three obvious points of similarity between the test and our measure of +/- cue use.

First, we have emphasized that the difficulty in our task concerns <u>initial</u> detection/selection of a goal. If the task has once been done correctly, it will not subsequently be failed. Again this relates to the idea that the CE is especially important in first setting up a novel performance, and that its role diminishes once a correct sequence of operations is practised. Similarly in a test like the Culture Fair, the whole point is that each new item requires a novel analysis. One would hardly call it a test of problem solving if each item involved exactly the same sources of stimulus variation requiring exactly the same type of solution.

Second, we have seen that explicit verbal prompts are one very strong cue controlling goal selection. Even subjects who failed after initial instruction were likely to pass with increased verbal emphasis, and everybody passed when this was increased to the level of an explicit commentary on errors and request for success. In a test like the Culture Fair, again, part of the point is that many requirements are left undescribed in the verbal instructions. Given the item in Figure 6, for example, one does <u>not</u> say to the subject, "Make sure that your choice satisfies variations in shape, in elongation and in shading"!

A third possible factor may be the <u>number</u> of concurrently-specified goals. In our task, use of the +/- cue is a third task requirement, specified after the first two; and similarly in the example Culture Fair item, it seems clear that a part of the difficulty arises in having to account for three separate sources of variation simultaneously (Carpenter et al., 1990). This last suggestion was the focus of Experiment 7, in which we increased the number of specified requirements by adding a concurrent task.

In this experiment, the letter detection task was exactly the same as before. Occasionally, dots also appeared above or below the stream of letters and digits coming up in the centre of the screen. Dots were presented for 200 msec, concurrently with a letter or digit pair. One dot appeared during the first part of each trial, concurrently with one of the symbol pairs 2 to 7; in addition, a quarter of the trials had a second dot, immediately after the +/- cue, i.e. concurrently with pair 11. In addition to performing the letter task as usual, subjects responded to the dots with a two-alternative speeded keypress response, with one key for dots above the symbol stream and another for dots below.

A critical manipulation was the order of instruction for the two tasks. As before, all instructions were given at the outset, before any practice trial. For half the subjects the letter task was described before the dot task, while for the other half, the dot task was described first.

We had three main predictions. First, if the number of concurrently-specified goals is important, the relationship between g and neglect of the +/- cue might become even stronger. Second, a similar relationship might appear for neglect of the dots, which though easily visible were briefly presented and spatially separated (about 3.5 degrees visual angle) from the central symbols. Third, if goal detection/selection is increasingly likely to fail as the number of specified goals increases, then results should depend on the order of instruction. Again we wished to confirm that even subjects who initially neglected part of the task had understood the instructions and were capable of complying with them. At the end of block 1 we asked subjects to repeat the rules both for use of the +/- cue and response to the dot, and in block 2 explicit feedback was given on cue mistakes or omission of dot responses. We tested 38 subject panel members between the ages of 40 and 50.

Again we replicated our findings on the relationship between g and cue neglect in block 1 (Figure 7). At .56 the correlation between these two was slightly higher than before - especially noting that the distribution of Culture Fair scores in this experiment (see bottom of figure) was extremely bunched around the mean. In this case, however, we observed no dependency on order of instruction, rate of failure and g correlation being much the same whether the letter task was described first or second. Again, all subjects except one repeated the +/- rule correctly when asked at the end of the block.

Insert Figure 7 about here

Still on the letter task, the hazard function for first correct pass is shown in the bottom row of Table 3. Results from block 1 were less clear than before, since with the concurrent task subjects picked the task up more slowly. The block 2 results, however, again confirmed that all subjects began to perform correctly with explicit verbal feedback.

The most interesting results concern omissions in the dot task, specifically in responses to the early dot, presented somewhere between positions 2 and 7 on every trial. (The occasional second dot presented immediately after the +/- cue was often missed throughout practice. The data will not be discussed.) Again the general pattern of performance was the same - a series of 0...n trials on which no response was made, with only rare omissions after the first correct response. When the dot task was described first, the mean number of omitted responses in block 1 was 3.17, and the correlation with Culture Fair errors was .12. But when the task was described second, the mean number of omissions rose to 6.61, with a correlation of .64. Again, all subjects described the rule correctly at the end of the block, and all began to respond to the dots when feedback was given in block 2.

In retrospect it is perhaps not surprising that the order of instruction was effective only for the dot task. Unlike the letter task, the dot task has rather simple requirements and can be explained in a few sentences. Adding these sentences before the complex letter task instructions has little effect on responses to the +/- cue, which in any case are described after a good deal of prior material. In contrast, responses to the dot task are very much influenced by moving the instructions to the end. Especially given the small number of subjects in each group, the results need to be replicated. As they stand, however, they do support the hypothesis that goal detection/selection by the CE runs into difficulty as the number of specified goals increases.

Summary

These experiments on mismatch between knowledge of a task's requirements and the corresponding behaviour seem very promising. At the simplest level, they confirm the link between frontal dysfunction and Spearman's g, people with low g scores showing a characteristic frontal error. Beyond this, the results suggest a particular functional role for the CE as a system responsible for recognising and/or selecting goals in a novel context. In all probability, such a system contributes rather little once appropriate, stereotyped behaviour has been established, or when the environment contains very strong cues (especially verbal) to goals/requirements. It is increasingly important, however, as the number of newly-specified goals increases.

Consistent practice and switches of set

As we have seen, a part of our strategy for analysing CE function is to investigate manipulations which alter a task's g correlation. For this purpose, it is useful to find tasks which, though relatively simple and well-specified, nevertheless have substantial g involvement. Our last experiments concern one such case, an "odd-man-out" reaction time (RT) task using simple geometric materials.

An example stimulus is shown in Figure 8. There are four panels, each containing a drawing of two cones. Drawings vary in 5 attributes: Shape of cone, shading of point, separation of cones, number of intersecting tick marks, and direction of pointing (cones pointing towards one another or both in the same direction). (A fifth attribute - horizontal or vertical arrangement of the two cones - is always irrelevant.) On each trial, one attribute is known to be relevant, and the task is to find the panel that differs from the others on this attribute. The response is made as quickly as possible on a corresponding 4-choice keyboard, with the keys laid out in a square. For all irrelevant attributes, there are always two panels with one value and two with another, i.e. no "odd-man-out" panel exists.

Insert Figure 8 about here

Preliminary data - described in our original proposal - were obtained from the same sample of 90 subjects used in Experiment 4. Each stimulus attribute was relevant for one block of trials, consisting of (usually) 2 practice trials followed by 8 experimental trials. Thus there were 5 blocks of trials using the stimuli shown in Figure 8, along with 5 more blocks using a different stimulus set constructed on similar principles. The correlation between mean RT (across the 10 blocks) and Culture Fair errors was .61. Even for a single block of only 8 responses, with a fixed relevant attribute, the average correlation was .49. These figures contrast with the usual finding that simple, well-specified RT tasks usually correlate with g no higher than about .30 (Hunt, 1980), as long as the g distribution is reasonably representative of the normal population. Indeed, the same 90 subjects were also given a standard same-different matching RT task, using pairs of upper case letters. Its correlation with Culture Fair errors (based on 40 trials) was .31.

Experiment 8 was designed with three hypotheses in mind. First, a key factor might simply be the complexity of the panels task. As compared with letter matching, it involves more stimuli which must be scanned using several fixations, more alternative responses etc. Second, it might be important that we gathered data only very early in practice. Third, the task required frequent changes of mental set. Each attribute was relevant for only 10 trials, and the rest of the time was to be disregarded. Again this relates to the idea that consistent practice with a fixed stimulus-response mapping diminishes CE involvement; perhaps frequent changes of set prevent the development of automatic performance. Indeed our task has substantial face similarity to various card-sorting tests requiring that the same stimulus materials be classified in different ways; difficulty changing set in such tasks is characteristic of frontal patients (Milner, 1963).

To address these three hypotheses, Experiment 8 used the following practice schedule. For the first 40 trials, the relevant attribute was fixed. For all subjects this attribute was separation of the cones. In a second set of 40 trials, a new relevant attribute was specified on each trial, each of the 5 attributes being used 8 times. Data were collected from the same group of 41 elderly adults tested in Experiment 6, though 4 were dropped because of failure to understand the instructions or experimenter error.

The following predictions may be derived from our three hypotheses. If the key factor is simply complexity, the g correlation (correlation between RT and Culture Fair errors) should be comparable to that in the previous experiment, and sustained throughout the 80 trials. If the key factor is practice, the correlation should decrease smoothly throughout these trials. If the key factor is set switching, the correlation should be low in the first 40 trials, but high in the second 40.

Results are shown in the upper panel of Figure 9. The initial set of 40 fixed-set trials has been divided into 8-trial blocks, labelled blocks 1 to 5. Results from the second set of trials are shown separately for each relevant attribute. Since the 8 trials for each attribute were distributed throughout these 40 varied-set trials, results have been plotted at the midpoint, block 8. Results for "cone separation" trials are shown as a circle, with the other 4 attributes as triangles.

Insert Figure 9 about here

The first point to note is that our prior results were replicated reasonably well in varied-set trials. Correlations with the Culture Fair for the individual attributes ranged from .32 to .50; for average RT across all 5 attributes the correlation was .52. In all probability, this is slightly lower than the value previously obtained only because of restriction of range of g in this elderly sample (see Figure 5).

Second, though the trend in g correlation across the 5 fixed-set blocks was not entirely clear, the tendency was for this correlation to begin at about the same level as seen in the varied-set trials, but to decline across blocks. The dotted line in the figure is the best linear fit to the data from blocks 1 to 5, extrapolated to block 8. Note though that this line is heavily dependent on the single low correlation obtained in block 5.

With this last caveat in mind, the data perhaps suggest the following picture. In fixed-set trials, the g correlation starts around .40 (for an 8-RT block) but declines with practice. Introduction of the varied-set procedure offsets this effect of practice, returning the correlation to approximately its original level, and at least to a level above the prediction we should obtain by extrapolating the practice curve derived from blocks 1 to 5.

In Experiment 9 we sought to replicate these results, and to investigate one particular aspect of the task's "complexity" - presence of irrelevant stimulus variation. In each display of Experiment 8, panels differed on all possible attributes, even though only the relevant attribute could be used to define an "odd man out". Does the need to ignore irrelevant stimulus differences contribute to g involvement? We used the following design. For the first 120 trials, the relevant attribute was always shape. These fixed-set trials were divided into three 40-trial blocks, with respectively 0, 2 and 5 irrelevantly-varying attributes per display (including horizontal-vertical arrangement of the cones). Thus block 1 reduced to a simple physical-match task, having three identical panels and one differing only on the relevant attribute; though it was still true that attributes other than shape varied across trials. A second set of 120 trials used the varied-set procedure, with a new relevant attribute specified on each trial. Again, there were three successive blocks of 40 trials each, with respectively 0, 2 or 5 irrelevantly-varying attributes per display. We tested the same 38 subjects used for Experiment 7.

Results, plotted in the same way as before, are shown in the bottom panel of Figure 9. Again the 40-trial fixed-set blocks have been divided into 8-trial sub-blocks, and a best-fitting straight line has been extrapolated from these data. This time, circles denote trials on which the relevant attribute was shape, with triangles for other attributes.

Results were similar to those from Experiment 8. Correlations with the Culture Fair test decreased systematically through the 120 trials of fixed-set testing, though much more slowly than before. With the introduction of variations in set, however, correlations returned to their initial level. In neither case did the number of irrelevantly-varying attributes have any apparent effect (though of course there was a substantial effect on absolute RT). As for overall correlation with g, results from the original experiment were replicated again, allowing for restriction of range. In the condition most comparable to the original experiment, varied set with irrelevant variation on all attributes (block 6), mean RT across the 40 trials correlated .53 with Culture Fair errors.

Given the noise in our data, it would be useful to replicate the findings using larger groups of subjects and more extended practice. As they stand, however, the results suggest that both practice and set switching are important factors in g involvement. The correlation reduces quite rapidly with consistent practice and a fixed mental set (Ackerman, 1988), but this effect is eliminated when variations in set are introduced. The results support the suggestion that consistent practice with a simple set of mental operations soon diminishes any input from the CE.

In further work we plan to look for boundary conditions on these results. In our original letter match data, for example, there was no suggestion of a change in g correlation over 50 trials of consistent practice. Instead the correlation (for 10-trial blocks) began around .25-.30 and stayed constant. We still wish to know what aspect of the panels task is responsible for the higher correlation seen at the start of practice, and the reduction with practice that subsequently takes place.

Future work

We have mentioned immediate plans for the next stage of each project. For random generation, we have shown that verbal and manual versions behave in much the same way. The next step is to investigate directly whether they involve the same CE system, by looking at interference when they are performed concurrently. The longer term goal is to test the hypothesis that, while the "slave systems" of working memory are modality- and content-specific, the CE is not. For the experiments on mismatch between knowledge and behaviour, we wish to extend our study of conditions that promote the effect, as a lead to conditions under which behaviour is strongly dependent on the CE's role in goal detection/selection. We are especially keen to explore further the similarities between conditions that promote neglect of a task requirement, and conditions in standard "problem-solving" tasks like the Culture Fair. For the experiments on practice and set switching in RT, we shall continue to investigate the aspects of RT tasks that control the initial g correlation, and the effects of different kinds of practice and transfer. The focus here is on conditions allowing transfer of control from the CE to "automatic" processing.

As we described in the original proposal, our intention as the project developed was to establish a collaboration with Dr P. Kyllonen's group at Brooks AFB in San Antonio. Such a collaboration would allow us to test the large numbers of subjects needed to assess precise correlations, for example in comparing different versions of a test, as well as encouraging feedback of our theoretical analysis of working memory for Air Force use. Our original proposal was to initiate this collaboration as profitable directions emerged from our work in Cambridge, and it seems clear that by now we have reached this stage. The projects on both neglect of a task requirement and set switching/RT have produced fairly definite hypotheses, to test which we need to measure the precise impact of different task manipulations on g correlations. With this in mind, the process of establishing a collaboration is now well under way. Initial discussions between Drs Baddeley and Kyllonen took place in San Antonio in February 1991; these have been followed up by further discussions in August (Duncan and Kyllonen, Chicago) and October (Baddeley and Kyllonen, Texas), and we are now considering how best to develop joint software for use of our tests in Dr. Kyllonen's laboratory.

In the first year we have done little linking dual task interference to frontal dysfunction or g. In particular we have yet to begin the proposed studies examining "profiles" of dual task decrement, frontal impairment and g correlation across a battery of different tasks. Here again the link with Dr Kyllonen's group is proving invaluable, since they have been able to provide us with results from a very large-scale study on intercorrelations and hence g correlations in a large, diverse test battery (standard tests from the kit of Ekstrom, French, Harman, & Derman, 1976). We intend to make these data the basis for the "profile" studies.

Publications

Baddeley, A. (in press) Working memory or working attention? In A. Baddeley and L. Weiskrantz (Eds.), Attention: Selection, awareness and control. A tribute to Donald Broadbent. Oxford: Oxford University Press.

Duncan, J. (in press) Selection of input and goal in the control of behaviour. In A. Baddeley and L. Weiskrantz (Eds.), Attention: Selection, awareness and control. A tribute to Donald Broadbent. Oxford: Oxford University Press.

Oral presentations

Duncan, J. et al. Mismatch between knowledge and behaviour: Frontal lobe dysfunction and Spearman's g. Experimental Psychology Society, Brighton, July 1991.

Duncan, J. The central executive component of working memory. Cognitive Science Society, Chicago, August 1991.

Duncan, J. et al. Goal selection: Studies of frontal lobe dysfunction, dual task interference and Spearman's g. British Neuropsychological Society, London, November 1991.

Consultation

May 1991, Austin, Texas. Baddeley - advice to Drs I. Foss and P. Kyllonen on studies in individual differences and language processing.

October 1991, Iowa City. Baddeley - advice on working memory and performance measurement at ONR contractors' meeting.

References

Ackerman, P.L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. Journal of Experimental Psychology: General, 117, 288-318.

Anderson, J.R. (1983). The architecture of cognition. Cambridge, Mass: Harvard University Press.

Baddeley, A.D. (1966). The capacity for generating information by randomization. Quarterly Journal of Experimental Psychology, 18, 119-129.

Baddeley, A.D. (1986). Working memory. Oxford: Oxford University Press.

Baddeley, A. (in press) Working memory or working attention? In A. Baddeley and L. Weiskrantz (Eds.), Attention: Selection, awareness and control. A tribute to Donald Broadbent. Oxford: Oxford University Press.

Baddeley, A.D., and Hitch, G. (1974). Working memory. In G.A. Bower (Ed.), Recent advances in learning and motivation, vol. 8. New York: Academic Press.

Brown, J.I., Nelson, M.J., and Denny, E.C. (1973). Examiner's manual. The Nelson-Denny reading test. Boston: Houghton-Miffin.

Bryan, W.L., and Harter, N. (1899). Studies on the telegraphic language. The acquisition of a hierarchy of habits. Psychological Review, 6, 345-375.

Carpenter, P.A., Just, M.A., and Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. Psychological Review, 97, 404-431.

Cattell, R.B. (1971). Abilities: Their structure, growth and action Boston: Houghton-Mifflin.

Duncan, J. (1990). Goal weighting and the choice of behaviour in a complex world. Ergonomics, 33, 1265-1279.

Duncan, J., Williams, P., Nimmo-Smith, M.I. and Brown, I. (in press). The control of skilled behaviour: Learning, intelligence, and distraction. In Attention and performance XIV, (ed. D. Meyer and S. Kornblum). Erlbaum, Hillsdale, N.J..

Duncker, K. (1945). On problem solving. Psychological Monographs, 58, (Whole No. 270, 1-113).

Ekstrom, R.B., French, J.W., Harmon, H.H., and Derman, D. (1976). ETS kit of factor-referenced cognitive tests. Princeton, N.J.: Educational Testing Service.

Frith, C.D., Friston, K., Liddle, P.F., and Frackowiak, R.S.J. (1991). Willed action and the prefrontal cortex in man: A study with PET. Proceedings of the Royal Society London B.

Hebb, D.O., and Penfield, W. (1940). Human behavior after extensive removal from the frontal lobes. Archives of Neurology and Psychiatry, 44, 421-438.

Hunt, E. (1980). Intelligence as an information-processing concept. British Journal of Psychology, 71, 449-474.

Institute for Personality and Ability Testing. (1959). Measuring intelligence with the culture fair tests. Champaign, Illinois: The Institute for Personality and Ability Testing.

Luria, A.R. (1966). Higher cortical functions in man. London: Tavistock.

Luria, A.R., and Tsvetkova, L.D. (1964). The programming of constructive ability in local brain injuries. Neuropsychologia, 2, 95-108.

Mettler, F.A. (Ed.) (1949). Selective partial ablation of the frontal cortex: A correlative study of its effects on human psychotic subjects. New York: Hoeber.

Milner, B. (1963). Effects of different brain lesions on card sorting. Archives of Neurology, 9, 90-100.

Nelson, H.E. (1982). National Adult Reading Test (NART). Test Manual. Windsor, U.K.: NFER-Nelson.

Norman, D.A., and Shallice, T. (1980). Attention to action: Willed and automatic control of behavior (Report No. 8006). San Diego: University of California, Center for Human Information Processing.

Raven, J.C., Court, J.H., and Raven, J. (1988). Mill Hill vocabulary scale. 1988 Revision. Oxford: Oxford Psychologists Press.

Schneider, W., and Shiffrin, R.M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. Psychological Review, 84, 1-66.

Spearman, C. (1927). The abilities of man. New York: Macmillan.

Weinstein, S., and Teuber, H.-L. (1957). Effects of penetrating brain injury on intelligence test scores. Science, 125, 1036-1037.

Table 1. Experiment 3. WAIS and Culture Fair IQs of 3 frontal patients.

Patient	Age	Frontal Lesion	Other Lesion	Current WAIS IQ	Culture Fair IQ
A	53	left	-	126	87
B	29	bilat	-	130	109
C	52	left	-	128	99

Table 1. Experiment 3. WAIS and Culture Fair IQs of 3 frontal patients.

Patient	Age	Frontal Lesion	Other Lesion	Current WAIS IQ	Culture Fair IQ
A	53	left	-	126	87
B	29	bilat	-	130	109
C	52	left	-	128	99

Table 2. Experiment 4. Distribution across subjects of the number of subblocks passed in block 1.

number passed	3	2	1	0
number of subjects	59	14	2	15

Table 3. Experiments 4, 6 and 7. Hazard function for first pass (probability of pass | no prior pass).

	Sub-block								
	1a	1b	1c	2a	2b	2c	3a	3b	3c
Experiment 4	.67	.47	.06	.40	.11	.00	.00	.13	.14
Experiment 6	.37	.23	.15	.29	.08	.00	.45	.50	1.00
Experiment 7	.37	.50	.33	.38	.60	1.00			

Table 4. Experiment 5. Data from 7 frontal patients.

Patient	Age	Frontal Lesion	Other Lesion	Control IQ[1]	Culture Fair IQ	+/- cue use Pass = 0 Fail = 1
1	36	right	-	118	96	0
2	39	bilat	-	98	87	1
3	21	bilat	-	94	84	1
4	55	bilat	-	98	91	1
5	37	bilat	petechial midbrain haemorrhages	102[2]	94	0
6	44	right	-	97	76	1
7	45	bilat	L temporal pole	94	97	1

1 Mean Culture Fair IQ of 2 or 3 control subjects matched to patient on age, sex, socioeconomic group, and score on National Adult Recoding Test (Nelson, 1982)

2 Incomplete data

Figure legends

Figure 1. Experiments 1 and 2. Per cent digram redundancy as a function of generation speed.

Figure 2. Activation of candidate goals. The current state activates possible next-states, while the goal state activates relevant next-states.

Figure 3. Example trial in the letter monitoring task. Time runs from top to bottom in the figure; all stimuli are actually presented in the centre of the screen. The initial instruction WATCH RIGHT is presented for 1 sec, and subsequent stimuli for 200 msec each with a 200 msec ISI.

Figure 4. Experiment 4. Proportion of subjects neglecting cue throughout block 1 (zero passed subblocks) as a function of Culture Fair score. Mean and standard deviation from the published norms (Institute for Personality and Ability Testing, 1959) have been used to transform Culture Fair IQs to z-scores. Note that extreme bins include all subjects beyond plus or minus 1.5 standard deviations from the mean.

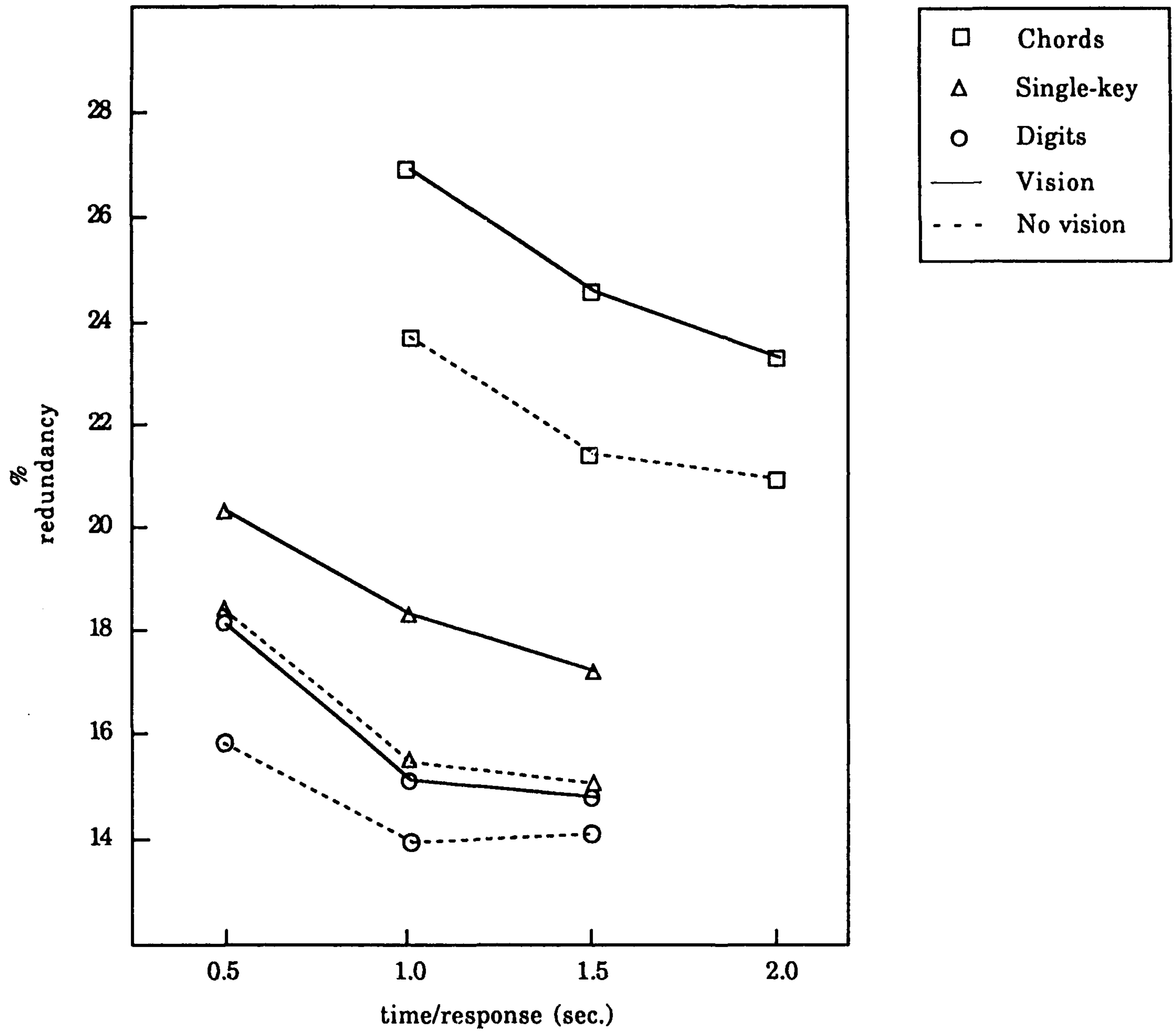
Figure 5. Experiment 6. As Figure 4, except that there were no subjects with Culture Fair scores above +0.5.

Figure 6. Example item designed to illustrate one kind of material in the Culture Fair test.

Figure 7. Experiment 7. As Figure 4.

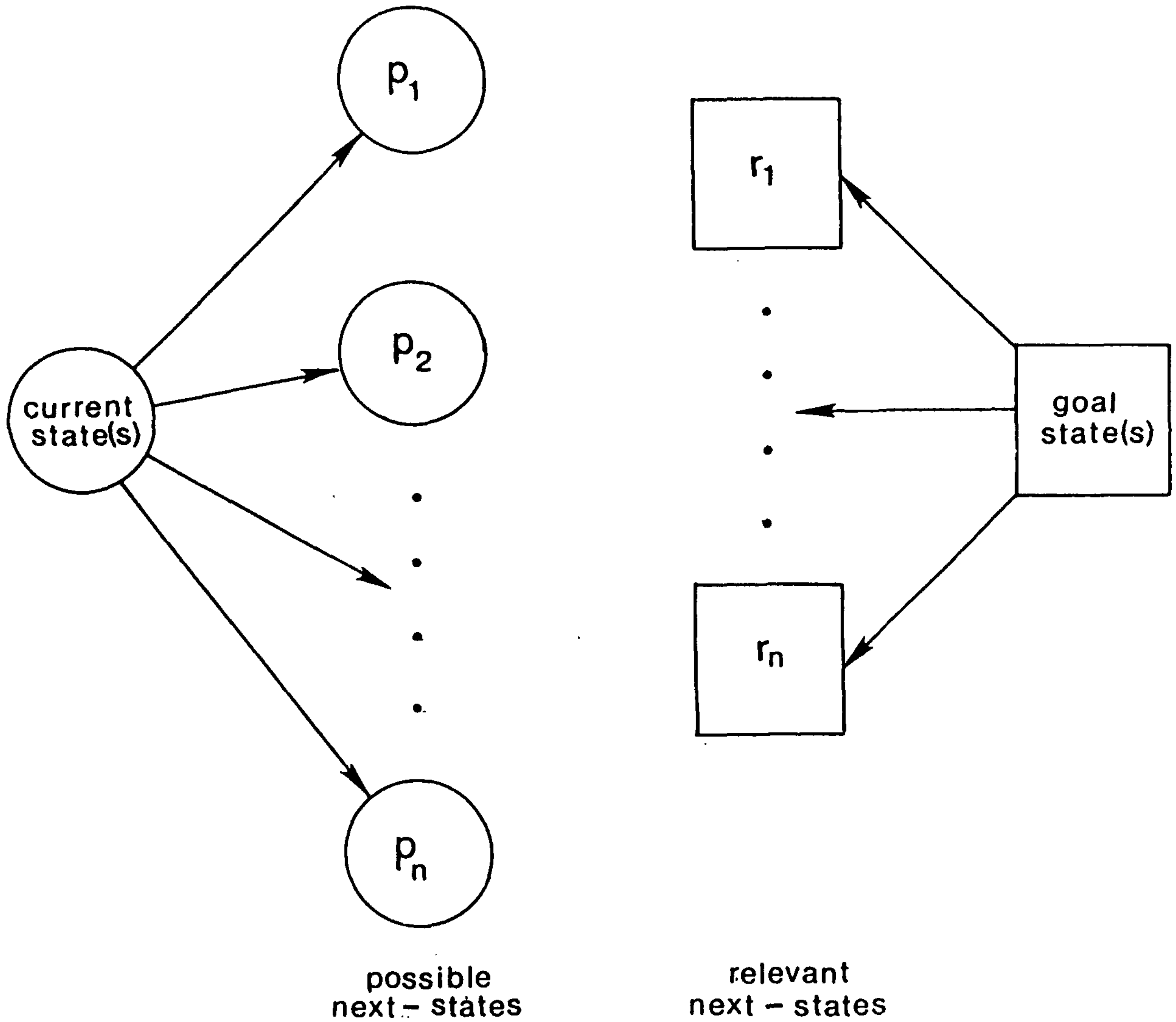
Figure 8. Sample display from the panels test.

Figure 9. Correlations between RT and Culture Fair errors. Top: Experiment 8. Bottom: Experiment 9.



% redundancy
28
26
24
22
20
18
16
14
0.5
1.0
1.5
2.0
time/response (sec.)
Chords
Single-key
Digits
Vision
No vision

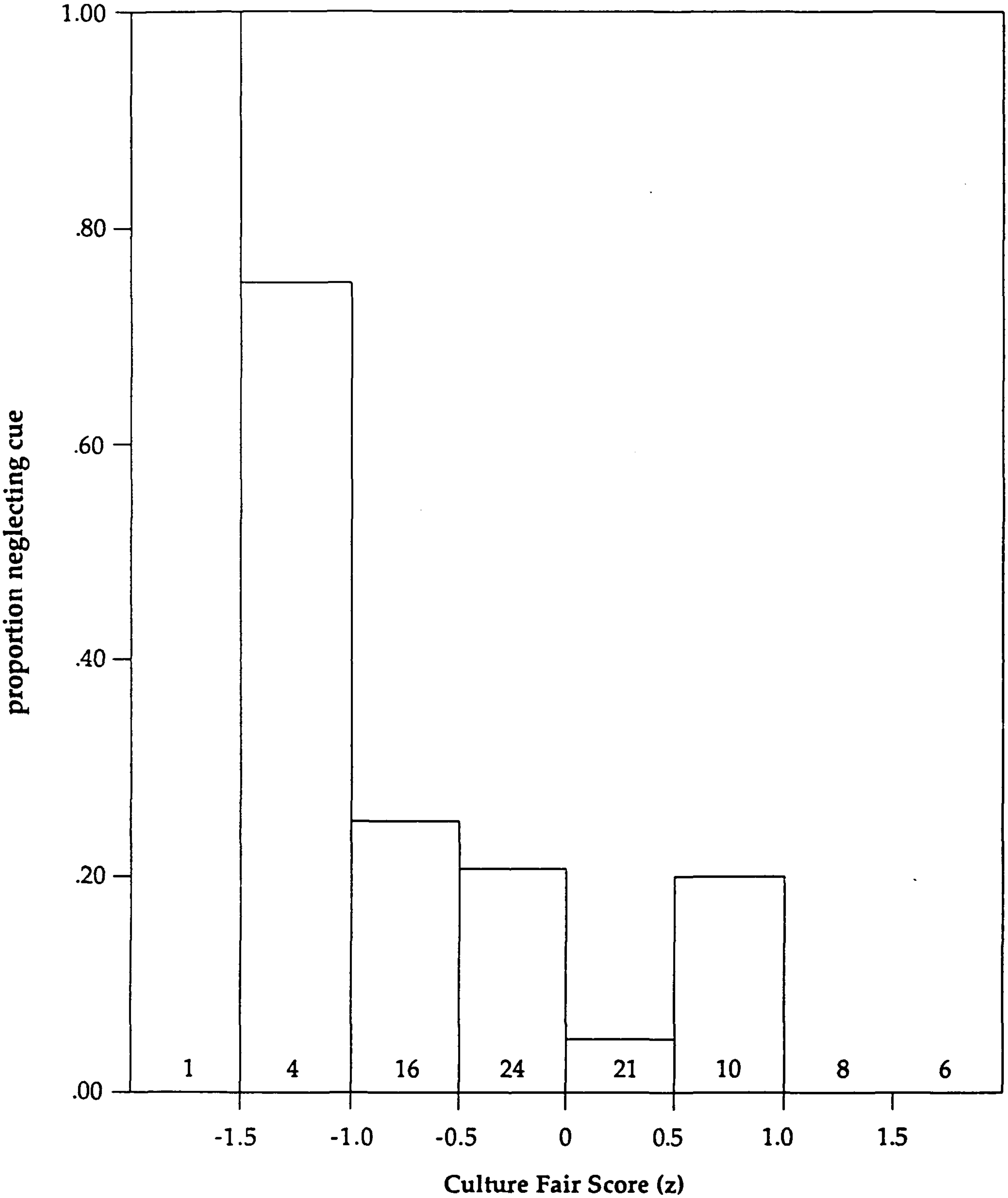
2


p1
p2
pn
current state(s)
r1
rn
goal state(s)
possible next - states
relevant next - states

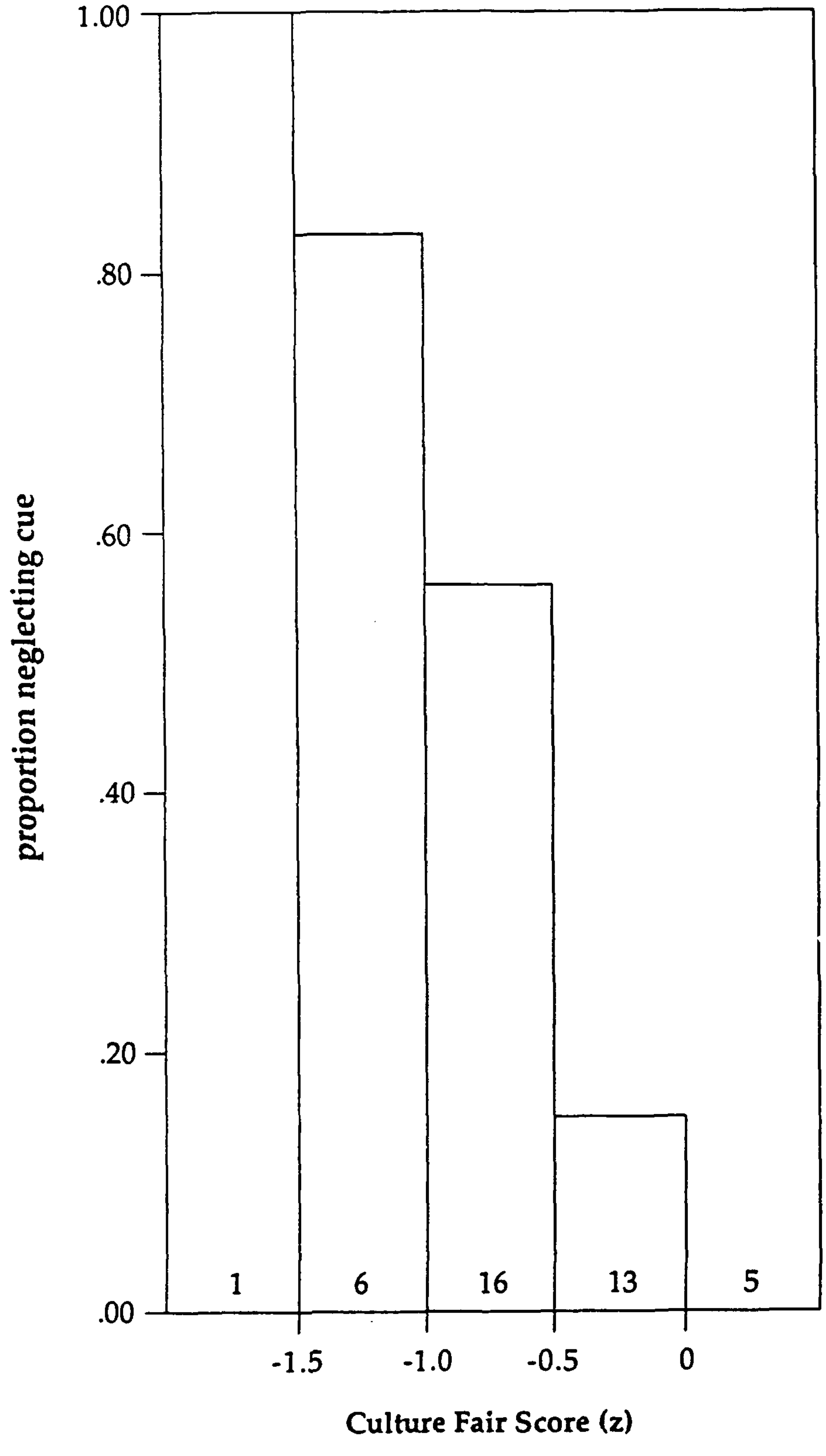
-3

WATCH RIGHT

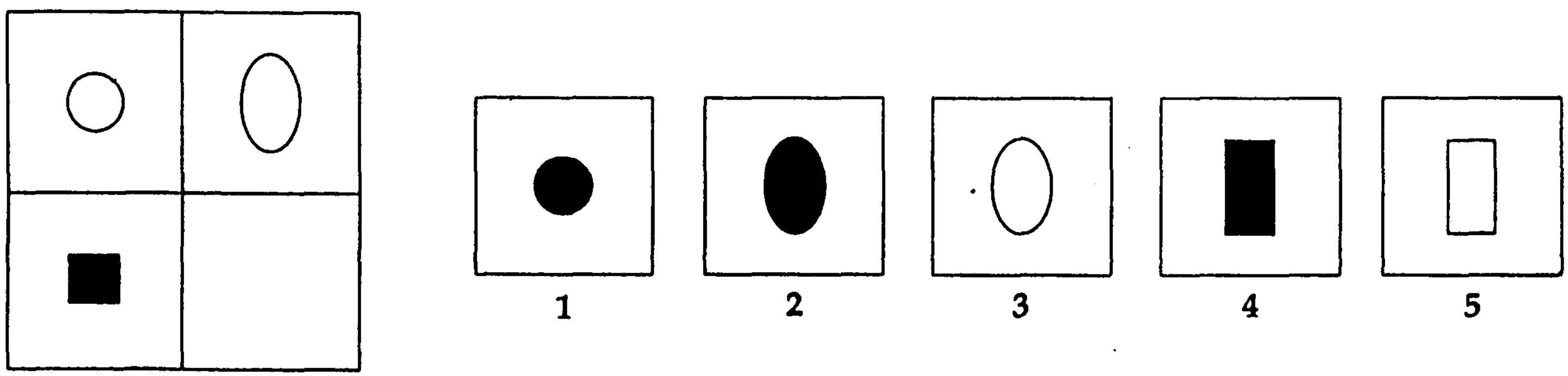
	2	3
	X	E
	B	C
	7	2
	4	4
	H	A
time ↓	L	Q
	5	9
	3	8
	T	M
	+	
	8	5
	N	F
	R	Y

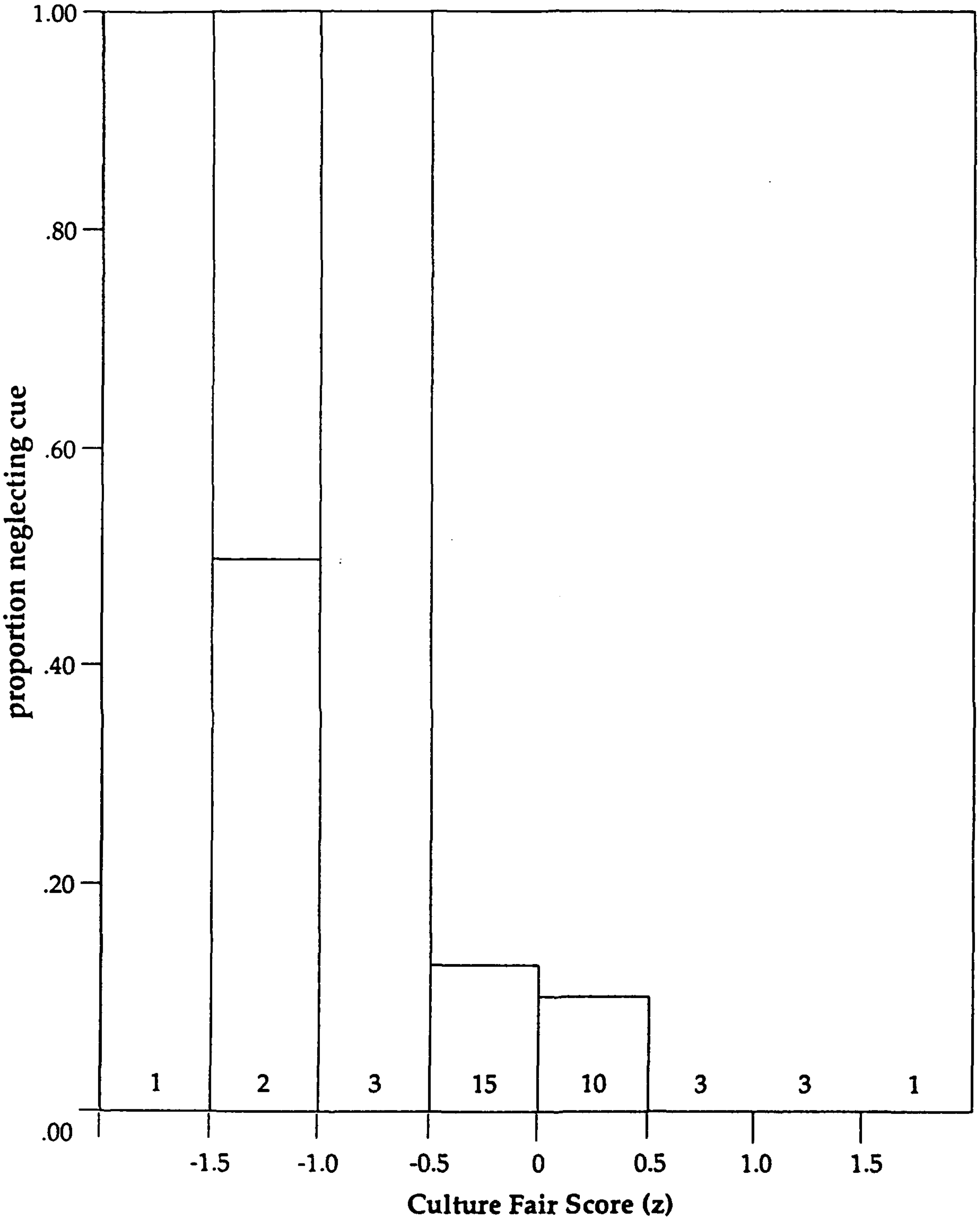

1.00
.80
.60
.40
.20
.00
proportion neglecting cue
1
4
16
24
21
10
8
6
-1.5
-1.0
-0.5
0
0.5
1.0
1.5
Culture Fair Score (z)

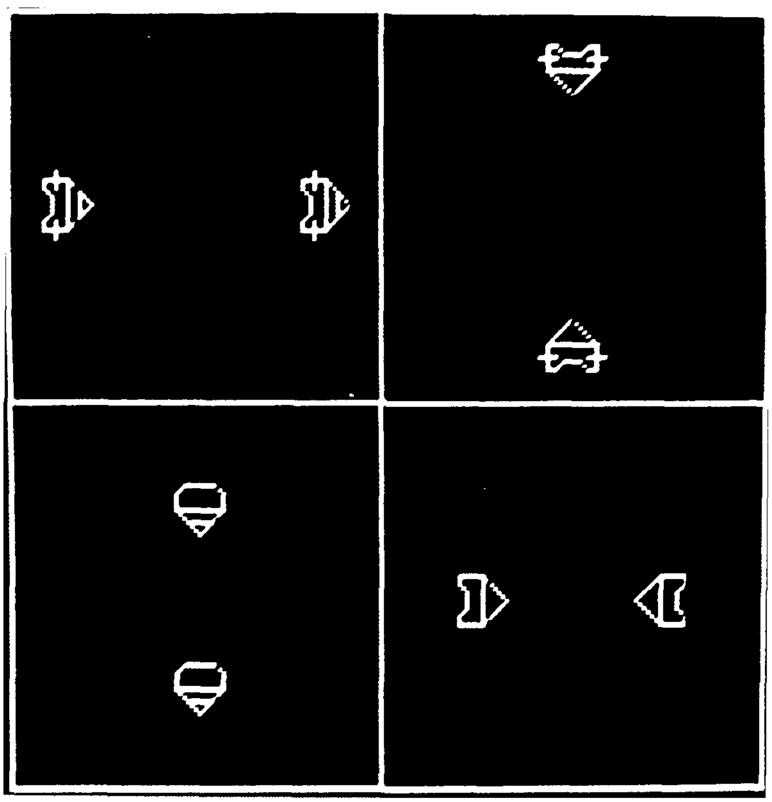
5



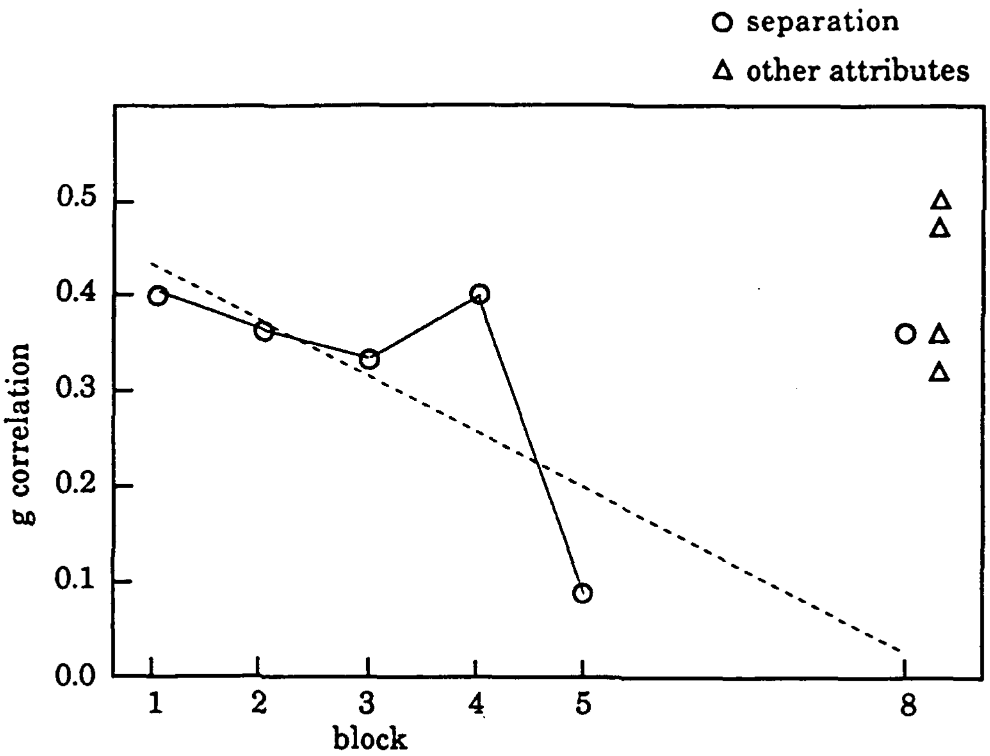
1.00
.80
.60
.40
.20
.00
proportion neglecting cue
1
6
16
13
5
-1.5
-1.0
-0.5
0
Culture Fair Score (z)


1
2
3
4
5

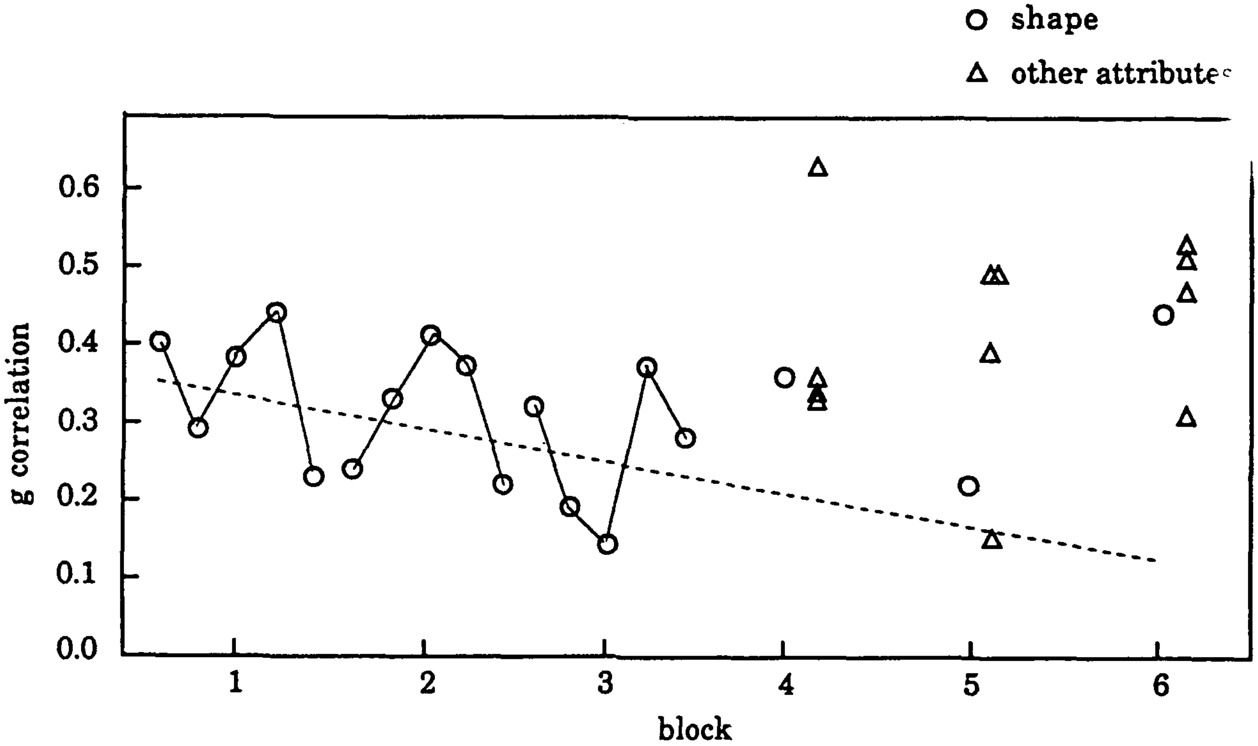

1.00
.80
.60
.40
.20
.00
proportion neglecting cue
1
2
3
15
10
3
3
1
-1.5
-1.0
-0.5
0
0.5
1.0
1.5
Culture Fair Score (z)



9



separation
other attributes
0.5
0.4
0.3
0.2
0.1
0.0
g correlation
1
2
3
4
5
8
block



shape
other attributes
0.6
0.5
0.4
0.3
0.2
0.1
0.0
g correlation
1
2
3
4
5
6
block