Neuropsychological verbal tests may lack “hold” properties in occupational studies of neurotoxic effects

Hans Michelsen, Ingvar Lundberg

Abstract

Objectives—In epidemiological studies of neurotoxic effects neuropsychological tests are often applied to assess possible functional changes. In these studies tests presumed to be resistant to neurotoxic effects, “hold tests”, are often used to adjust performance in tests presumed to be sensitive to neurotoxic effects, “non-hold tests”, in assessment of pre-exposure ability. This conception is based on experience from the examination of patients with organic brain disorders. For this adjustment to be valid, hold tests should change comparatively little over time and should not change differently in groups of people with different exposures to neurotoxic agents.

Methods—These assumptions were examined in 71 carpenters and 135 painters divided in three subgroups according to level of cumulative exposure to organic solvents. The results were noted from two verbal tests (test A, which involved following verbal instructions and test C, word comprehension), which were performed at conscription (age 18–20). The tests resemble hold tests used in occupational studies. The same tests were performed again at the age of 45–60 together with a conventional synonym test (SRB 1), often used in occupational studies.

Results—In the three tests given at the time of the investigation some differences were found between the carpenters and the subgroups of painters. The painters with low exposure tended to perform better than the carpenters in all three tests and the heavily exposed painters tended to perform less well in the tests. These differences were not present at the age of 18–20. In one of the conscription tests painters with different cumulative exposure to organic solvents developed differently over time. There was little improvement among painters with heavy and intermediate exposure and obvious improvement among painters with low exposure. The results in the other conscription test showed similar tendencies.

Conclusions—The results indicate that the hold tests examined do not meet basic criteria for hold tests used in occupational studies. Thus adjustments for conventional verbal tests seem inappropriate in such studies. It is possible that other presumed hold tests may meet the criteria but our results suggest that such tests also should be evaluated before being used as hold tests in working populations.

Keywords: conscription tests; longitudinal study; organic solvent exposure

To assess the possible neurotoxic influence of long-term exposure to organic solvents, and other neurotoxic agents, various neuropsychological tests have been used. Such tests have been applied widely when examining groups of patients to find a possible functional change—for example, in relation to a suspected early dementia or a psycho-organic syndrome according to the diagnostic system in psychiatry DSM III. One problem of these examinations is to establish the subject's original level of intellectual capacity, with which his or her current functional level must be compared. Several methods have been applied to estimate previous or primary intellectual levels. In a few cases test results from early periods in the person's life may be obtained—for example, tests performed in connection with military service. These types of data, mirroring the person's functional level almost always before occupational exposure, are desirable but generally difficult to get access to.

To use only current test data to measure a functional change is standard practice in neuropsychological assessment of organic brain disorders. The fundamental assumption is that certain cognitive functions are more stable, whereas others are more sensitive and more easily affected by the disease under study. For a long time the type of “hold test” used to best assess general primary intellectual level has been a verbal test such as a vocabulary test.

In one commonly used test battery, WAIS R, several single subtests such as information, vocabulary, comprehension, certain indices of these, and other verbal subtests have also been used as hold tests. This procedure is based on observations that many patients showing various organic changes retain well established, often over learned, verbal skills longer than other intellectual functions such as short-term memory, problem solving ability, or perceptual ability.

When, in the 1970s, the neuropsychiatric effects of exposure to organic solvents started to be investigated in Sweden it was found that symptoms experienced by exposed workers were similar to those of patients with the early stages of diffuse organic mental disorders.
Thus it seemed natural to apply psychometric tests as outcomes when trying to corroborate the putative effects. A test battery was compiled (TUFF battery\(^{11,12}\)), which has since been used at the departments of occupational medicine in Sweden. The TUFF battery contains subtests of verbal ability and is similar to the general recommendations of the World Health Organisation (WHO\(^{13}\)) on the construction of test batteries in connection with neurotoxic effects.

In many studies of neuropsychiatric effects of exposure to solvents verbal tests were used as an indicator of pre-exposure performance, either for adjustment of non-hold test results or simply as a background for the discussion of results in non-hold tests.\(^{10,14-18}\) Sometimes adjustment for hold test results have notably changed the results in non-hold tests obtained without such adjustment.\(^{19}\)

Although verbal hold tests are obviously valuable to establish premorbid intellectual capacity when examining patients with signs of early dementia or psycho-organic syndromes,\(^{20}\) it seems not to have been investigated whether verbal hold tests are actually appropriate for the same purpose when groups of workers are examined for less serious forms of functional changes—for example, the neurotoxic effects expected among workers with long term exposure to organic solvents. With the clinical research model of comparing hold test results with non-hold test results in occupational epidemiology it is assumed that on average hold tests will show the level of performance which would have been obtained through non-hold tests if there had been no neurotoxic effects. If this is valid hold tests must firstly assess the intellectual level as it was before neurotoxic effects deteriorated the subject’s performance in certain respects. Over periods of 10 years or more, certain changes in performance can be expected because of age, learning experiences, and other social and cultural factors. In verbal tests these changes may be positive, whereas in perceptual tests they may develop negatively.\(^{21,22}\) Therefore, test results are often expressed in points, standardised for age. However, it is necessary that hold tests are reasonably stable over time and that they change considerably less than non-hold tests. Secondly, hold tests must also develop similarly over time in socially and culturally similar occupational groups, or in groups within the same occupation but with different degrees of neurotoxic exposure.

The aim of the present study is to investigate whether some different measures of primary intellectual level meet the second and, to some extent, the first of these criteria, and may thus be used as hold tests in occupational studies of neurotoxic effects.

**Material and methods**

The group considered for the study consisted of the male painters and carpenters who were born in 1925 or later, who were members of the Stockholm chapters of their trade unions in 1965, and had belonged to the union for at least 10 years before 1971. This restriction was introduced to make sure that the painters were active during the 1950s and 60s, when solvent exposure in the painting trade was high.\(^{23}\) Data on the basic criteria were collected through the membership files of the respective unions, in which 767 painters and 1212 carpenters met the criteria mentioned.

To obtain measures of primary intellectual level, we used results from psychometric tests done in connection with conscription for military service at the age of 18–20. Over the years, the subtests making up the test batteries have varied. The study subjects were taken from five different test years, in which two subtests were identical. From each test year (1949, 1951, 1953, 1957, and 1961) all painters and a random sample of about half of the carpenters were selected, amounting to a total of 148 painters and 85 carpenters. They were invited to a clinical examination of the neuropsychiatric effects of exposure to solvents. Altogether 135 painters and 71 carpenters participated in the investigation, which was carried out in 1988–9. Details on the selection of the study group have been given elsewhere.\(^{24}\)

**PRIMARY INTELLECTUAL LEVEL**

Since the early 1940s all 18–20 year old men have been given a psychometric test battery in conjunction with conscription for compulsory military service. The results of these tests have been used, together with other data collected at the conscription, for assigning the conscripts to different types of military training and duties. The test battery has contained three to five tests mirroring certain aspects of general intellectual functions. The tests were of the paper and pencil type and were given as group tests according to a given manual.

The two tests which were identical in the test battery given in the five years selected for the investigation were one instruction test and one vocabulary test. Both tests measure verbal ability and overlap widely with general mental ability. In the first test (A), the subject is asked to follow simple instructions such as put a cross in a square or underline the fifth figure in a row of figures. The test involves simple verbal comprehension, discriminatory ability, simple problem solving, and for certain items, a certain amount of short term memory, in that the subject is asked to answer complex instructions. The test consists of 40 items to be answered in a maximum of 12 minutes. The other test (C) is a word comprehension test, in which each item contains five words. Four words are features of one common category, whereas the fifth word deviates. The fifth word is to be identified and is, thus, the correct answer. The test contains 40 such items of five words and is to be answered in a maximum of seven minutes. The results of these conscription tests, which the subjects took at the age of 18–20 (A\(_a\) and C\(_a\)) were obtained from the military archives (Krigsarkivet). In the entire groups from each test year the test result distributions were similar for painters and carpenters.
Table 1. Distribution of background variables that were notably associated with any of the psychometric test results, for occupational groups and for three levels of cumulative exposure

<table>
<thead>
<tr>
<th>Variable</th>
<th>Carpenters (n = 71)</th>
<th>Painters (n = 135)</th>
<th>Low (n = 34)</th>
<th>Intermediate (n = 67)</th>
<th>High (n = 34)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age ≥ 50</td>
<td>31 (44)</td>
<td>72 (53)</td>
<td>7 (21)</td>
<td>40 (60)</td>
<td>25 (74)</td>
</tr>
<tr>
<td>Alcohol consumption: Low</td>
<td>35 (49)</td>
<td>66 (49)</td>
<td>15 (44)</td>
<td>32 (48)</td>
<td>19 (56)</td>
</tr>
<tr>
<td>Intermediate</td>
<td>19 (27)</td>
<td>36 (27)</td>
<td>10 (29)</td>
<td>17 (25)</td>
<td>9 (26)</td>
</tr>
<tr>
<td>High</td>
<td>17 (24)</td>
<td>33 (24)</td>
<td>9 (26)</td>
<td>18 (27)</td>
<td>6 (18)</td>
</tr>
<tr>
<td>Dyslexia</td>
<td>19 (27)</td>
<td>28 (21)</td>
<td>10 (29)</td>
<td>12 (18)</td>
<td>6 (18)</td>
</tr>
<tr>
<td>Schooling &gt; 9 y</td>
<td>10 (14)</td>
<td>16 (12)</td>
<td>5 (15)</td>
<td>7 (10)</td>
<td>4 (12)</td>
</tr>
<tr>
<td>Vocational training in:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Vocational school</td>
<td>25 (35)</td>
<td>62 (46)</td>
<td>13 (38)</td>
<td>31 (46)</td>
<td>18 (53)</td>
</tr>
<tr>
<td>2. Apprentice</td>
<td>46 (65)</td>
<td>71 (33)</td>
<td>20 (59)</td>
<td>35 (52)</td>
<td>16 (47)</td>
</tr>
</tbody>
</table>

The two tests mentioned were also given to the subjects of the current investigation (A_{row} and C_{row}). The same instructions and marking norms as on the day of conscription were used. The tests were given as group tests to two subjects at a time as the first task in a one day examination after certain medical tests.

Another estimate of the subjects’ primary intellectual level was made through a verbal synonym test (SRB 1), which is part of the TUFF battery. This test is similar to the test C and contains 30 items in which a synonym for the stimulus word is to be selected from among five words. The testing time is seven minutes. The test was given individually by a psychologist as the first test in a major test battery.

Also, two indices of change in the two military conscription tests were devised, and made up of the difference (A_{off} and C_{off}) between the performance at the current investigation and that of the age of 18–20 (A_{off} = A_{row} - A_{10} and C_{off} = C_{row} - C_{10}).

Table 2. Correlation coefficients (r) between psychometric test results of painters and carpenters

<table>
<thead>
<tr>
<th>Tests examined</th>
<th>Carpenters (r)</th>
<th>Painters (r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Different tests on the same occasion:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A_{row}/C_{row}</td>
<td>0.56</td>
<td>0.59</td>
</tr>
<tr>
<td>A_{row}/C_{SRB}</td>
<td>0.57</td>
<td>0.63</td>
</tr>
<tr>
<td>A_{row}/C_{SRB}</td>
<td>0.69</td>
<td>0.49</td>
</tr>
<tr>
<td>C_{row}/SRB 1</td>
<td>0.79</td>
<td>0.47</td>
</tr>
<tr>
<td>Same test on different occasions:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A_{row}/A_{row}</td>
<td>0.68</td>
<td>0.63</td>
</tr>
<tr>
<td>C_{row}/C_{row}</td>
<td>0.62</td>
<td>0.57</td>
</tr>
<tr>
<td>Different tests on different occasions:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A_{row}/SRB 1</td>
<td>0.51</td>
<td>0.46</td>
</tr>
<tr>
<td>C_{row}/SRB 1</td>
<td>0.47</td>
<td>0.53</td>
</tr>
</tbody>
</table>

OCCUPATIONAL EXPOSURE

An occupational hygienist carried out extensive individual interviews with all subjects to survey the lifetime exposure to organic solvents. This information was summed up in a cumulative exposure index. The construction of the index has been described elsewhere. The 25% of the painters who had the highest cumulative exposure were categorised as the heavy exposure group, the 25% with the lowest cumulative exposure were the low exposure group, and the remaining 50% as the intermediate exposure group. All the carpenters had none or a very low cumulative exposure.

SOCIAL AND MEDICAL DATA

Certain background data, relevant to the assessment of the test results and primary intellectual level, were gathered from structured interviews in connection with the examination. The variables included dyslexia—that is, reported dyslexia in school or not, handedness (unambiguously right handed or not), schooling (equal to or more than nine years in school or not), and vocational training as an apprentice or in vocational school. A measure of long term alcohol consumption was established as described elsewhere (table 1).

Other questions dealt with neurological and other diseases and the use of drugs affecting the central nervous system but these factors were not used in the data analysis as they were not found to affect the test results under study.

STATISTICAL METHODS

Statistical analyses were performed for the groups of painters and carpenters, on the one hand, and on the other, for the group of painters divided into three subgroups with reference to cumulative exposure. Descriptive measures, mean (SDs), and Pearson product-moment correlation were calculated for all test variables for all groups of participants.

Multiple linear regressions were performed between the different solvent exposures and background variables and the outcomes of the respective psychometric tests and test indices. Age and alcohol consumption were included in all calculations, except in the analyses of test data from age 18–20. Other background variables judged to be able to affect the outcome such as schooling, vocational training, dyslexia, and medical factors of importance to the central nervous system, have been tested individually. In the final analysis, those factors were included for which the P value for the regression coefficient was < 0.1.

To study more notable effects of solvent on changes in the conscription tests, A_{row} and C_{off} were dichotomised at the 10th percentile of the distribution among carpenters, and logistic regression analyses were performed. In the A test the 10th percentile corresponded to currently worse performance of a least two units. Background variables investigated in these analyses were the same as those tested in the multiple regression analyses. The computer program EGRET was used for these analyses. For all other calculations the SAS computer program has been used.
Results

The correlations between the results of the five tests of primary intellectual level were calculated per occupational group (table 2). The agreement between different tests given at the same point in time varied between r 0.47 and 0.79. The highest correlation was obtained among carpenters in the two word comprehension tests, C1ow and SRB 1. The correlations for the repeated comprehension tests were between r 0.57 and 0.68. Generally lower correlations were obtained between different tests from different points in time (r 0.46–0.57). The correlations were slightly lower in the group of painters in six of the eight comparisons.

The occupational groups painters and carpenters performed more or less equally well on the two comprehension tests (A18 and C18) at age 18–20, on the repetition of the comprehension tests (A1ow and C1ow), and the verbal test SRB 1. The SDs were also very similar for each of the tests (table 3).

When comparing the results of the comprehension tests at age 18–20 with the corresponding current results, in the form of individual difference measures (A1ot and C1ot) similar improvements in the performance of two occupational groups were found (tests on the mean improvements in A and C in both groups resulted in P < 0.0001).

Comparing test performance between the differently exposed subgroups of painters with the group of carpenters, there was a tendency for painters exposed at a low level to perform better on average than the carpenters in all tests (table 3).

In the multiple regression analysis of A1ow, with adjustment for background variables there was a clear tendency for intermediate and heavily exposed painters to perform worse than carpenters (table 4). Among the controlled variables a higher A1ow and more schooling were associated with a better result in A1ow.

In the corresponding analysis of C1ow, a weak association remained for less exposed painters to perform better than carpenters and a higher C1ow and more schooling to be associated with better results.

In the regression analysis with SRB 1 less exposed painters performed better than carpenters. A better result in SRB 1 was strongly associated with a better result in A1ow and C1ow but dyslexia impaired the result in SRB 1.

In A1ot and C1ot painters with low exposure tended to improve more than carpenters, and intermediate and heavily exposed painters to improve less (table 3). The regression analyses between exposure and A1ot and C1ot with adjustment for appropriate background variables confirmed these tendencies for A1ot among intermediately and heavily exposed painters (table 4).

Painters with low exposure tended to obtain better results than intermediate and heavily exposed painters in all tests performed (table 3). The exposure-response trend among painters also remained after controlling for background variables (table 4), and was notable in A1ow among both immediately and heavily exposed painters. Clear differences were also obtained between less exposed and more exposed painters in A1ot and between less and immediately exposed painters in A1ot.

The logistic regression analyses of A1ot and C1ot gave results that were similar to those obtained by multiple linear regression. There were obvious differences in A1ot between the
heavily exposed painters and the carpenters, as well as between the painters with low cumulative exposure and painters with intermediate or high cumulative exposure (table 5). In the C_{int} variable there were no differences between carpenters and painters or between the painters' exposure categories.

Discussion

The results of this study exemplify the difficulty of finding stable measures of primary intellectual capacity to be used as hold tests. The tests included two from the age of 18–20 (A and C) repeated at the time of the current investigation and one further verbal test (SRB 1). They resemble the types of verbal tests often used as hold tests. Carpenters showed higher correlations between the different tests given at the current investigation than painters. The highest correlation was found between C_{row} and SRB 1, the two word comprehension tests.

In the three tests that were given at the time of the investigation, tendencies towards differences were found between the carpenters and the three exposure groups of painters after adjustment for appropriate background factors. The painters with low exposure were better in SRB 1 and C_{row} than the carpenters, and the intermediate and heavily exposed painters tended to perform less well than the carpenters in A_{row}.

A clear improvement over time was found in the conscription tests (A_{int} and C_{int}) for both occupational groups. The results show an exposure-response relation in the amount of improvement in one of the tests (A_{int}), the improvement being lower for those with higher lifetime exposure. Considering the similarity of the exposure-response associations among painters for SRB 1, A_{row}, and C_{row}, it may be reasonable to assume that a similar development over time had occurred for the SRB 1 as for the A and C tests.

Our findings indicate a possible association between exposure to organic solvents and the test results. It is possible that subjects who had various complaints and simultaneous deterioration in psychometric test results may have exaggerated their level of exposure. This could have resulted in a spurious exposure-response relation among painters. However, we think that we have not reduced the possible misclassification of exposure through the detailed exposure interview performed by an occupational hygienist.

Otherwise, potential causes of misclassification of either exposure or outcome—such as, a lack of motivation on the part of the subject—should most likely have been non-differentially distributed over exposure or outcome categories. Thus they would tend to diminish a true exposure-response relation.

Analyses were performed to study whether workers who had changed jobs showed different test results than those who did not. It was found that those who subsequently changed jobs, among painters as well as carpenters, had somewhat better results at the age 18–20 than those who were not to change. However, a change of job seemed not to affect the results found for solvent exposure in the regression analyses in neither the change in test results over time nor the test results at the re-examination if allowance was made for the test results at age 18–20.

Earlier studies on changes of test results and problem solving behaviour after lengthy periods of time indicate that they may be due to training and other kinds of life experiences. 21, 22 This may to some extent explain the improvements in the repeated conscription tests reported here. Another explanation may be that the subjects were better motivated to obtain a good result at the re-examination. However, none of these explanations can account for the differences in test outcomes between the exposure groups of painters.

Thus we think that the different development of test results over time among painters with different levels of cumulative solvent exposure is real. In our view two main explanations may be offered for this finding.

Firstly, exposed painters with a higher capacity for intellectual development might increasingly try to avoid work involving exposure to solvents whereas painters with a lower capacity for intellectual development would more commonly remain in this kind of work. This hypothesis of a selection among the painters is supported by the fact that carpenters and the whole group of painters showed similar mean (SD) test results at the age of 18–20 as well as at the re-examination.

Secondly, exposure to solvents had an effect on test performance. In this case one must assume that for social and cultural reasons linked to occupation painters and carpenters are equally likely to develop over time in terms of verbal abilities. This development is not affected by a low level of solvent exposure as found in the low exposure group in our study, whereas intermediate and heavy solvent exposure hamper this development in an exposure related manner. In the light of this explanation the tests used would in fact function as non-hold tests.

Our data are insufficient to examine the correctness of these two possible explanations. It may be concluded that the putative hold tests examined in this study do not seem to meet basic criteria for hold tests to be used in occupational studies of neurotoxic effects. However, the choice of the two conscription tests was restricted to those three to five tests that have been used in the actual study group when the subjects were 18 years of age. The two verbal tests selected were the only two conscription tests that were available for the entire study group and they have much in common with some of the subtests in WAIS-R, 4 often used as hold tests. One of the conscription tests also had a good correlation with the third test in this study which is the kind of verbal hold test used in similar studies. The two conscription tests seem to change considerably and differently over time in groups with different exposure to organic solvents and the results in the third test were also
different in groups with different exposure to organic solvents. These differences may not be explained by dissimilar distributions of other known determinants of test results in the exposure groups. Thus the use of such hold tests seems inappropriate if occupational exposure is expected to result in minor functional changes although more research should certainly be encouraged on this point. This, in turn, emphasises the need for a careful and thorough epidemiological study design to make investigated groups with and without exposure to neurotoxic agents comparable for primary intellectual capacity. More research is also needed to prove that other, and possibly more valid, hold tests are actually useful in studies of neurotoxic effects in working populations.

We thank Gun Nise who performed the exposure assessment, Annsika Gustavsson who performed the computer work, and Bengt Sjögren who gave valuable comments on the manuscript. The study was financed by the Swedish Work Environment Fund.

Neuropsychological verbal tests may lack "hold" properties in occupational studies of neurotoxic effects.
H Michelsen and I Lundberg

*Occup Environ Med* 1996 53: 478-483
doi: 10.1136/oem.53.7.478

Updated information and services can be found at:
http://oem.bmj.com/content/53/7/478

*These include:*

**Email alerting service**
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/