British Journal of Industrial Medicine 1989;46:483–495

Acute neurobehavioural effects of toluene

DIANA ECHEVERRIA, L FINE, G LANGOLF, A SCHORK, C SAMPAIO

From the Department of Environmental and Industrial Health, School of Public Health, University of Michigan, Ann Arbor, MI 48109, USA

ABSTRACT An acute inhalation chamber study of 42 college students was performed to investigate the relation between exposure to 0, 75, and 150 ppm of toluene and changes in central nervous system function and symptoms. Paid subjects were exposed for seven hours over three days. Verbal and visual short term memory (Sternberg, digit span, Benton, pattern memory); perception (pattern recognition); psychomotor skill (simple reaction time, continuous performance, digit symbol, hand-eye coordination, finger tapping, and critical tracking); manual dexterity (one hole); mood (profile of mood scales (POMS)); fatigue (fatigue checklist); and verbal ability were evaluated at 0800, 1200, and 1600 hours. Voluntary symptoms and observations of sleep were collected daily. An analysis of variance and test for trend was performed on the difference and score for each concentration reflecting an eight hour workday where each subject was their own control. A 3 × 3 Latin square study design evaluated toluene effects simultaneously, controlling for learning across the three days and the solvent order. Intersubject variation in solvent uptake was monitored in breath and urine. A 5–10% decrement in performance was considered significant if it was consistent with a linear trend at $p < 0.05$. Adverse performance at 150 ppm toluene was found at 6·0% for digit span, 12·1% for pattern recognition (latency), 5·0% for pattern memory (number correct), 6·5% for one hole, and 3·0% for critical tracking. The number of headaches and eye irritation also increased in a dose response manner. The greatest effect was found for an increasing number of observations of sleep. Overall, no clear pattern of neurobehavioural effects was found consistent with the type 1 central nervous system as classified by the World Health Organisation. Subtle acute effects, however, were found just below and above the ACGIH TLV of 100 ppm toluene, supporting the position that the guideline be lowered since the biological threshold of behavioural effects may be comparable with the TLV.

The acute behavioural effects of inhaling toluene after a single exposure are reversible,[1] but although reversible, they may also be an early sign of central nervous system (CNS) impairment leading to irreversible losses in performance with repeated exposures.[2,3] Acute decrements may be large enough to decrease the safety margin of industrial tasks or to lower productivity. Evidence from solvent studies suggests that symptoms might occur before other measurable acute CNS effects are apparent.[4–6]

The effects of toluene were first studied under controlled eight hour exposures in 1942.[4] Mild CNS symptoms and irritation were reported at 200 ppm, fatigue and muscular weakness were noted at 300 ppm, and headache, dizziness, and staggering were incapacitating at 600 ppm. Mild physiological effects have also been reported at 200 ppm.[7,8]

Later studies emphasised detecting changes in vigilance and visuomotor ability measured by psychomotor tests.[3,9–12] A decrement was reported for the correct response rate on a dual task involving isual vigilance and tone detection.[12] Two studies have enlarged the number of performance functions to include higher cortical function, visual perception, vigilance, visuomotor ability, manual dexterity, and simulations of industrial tasks.[13,14] Manual dexterity and visual perception were impaired, as measured by the response time on a peg board test and the Landolt's ring test. The effects of toluene have also been studied in the presence of p-xylene,[5] and ethanol[6]; no adverse behavioural effects were attributable to toluene, although CNS symptoms were reported.

The evidence suggests that inhaling toluene at 100 ppm over seven hours can cause slight impairments in vigilance, manual dexterity, and, perhaps, in visual memory. The objective of the present study was to determine whether these acute effects would be detected on a battery of tests used in epidemio-

Accepted 18 April 1988

logical studies[15-18] and to examine a wider range of human performance functions. This is the rationale for documenting acute effects at the industrially relevant levels of 75 and 150 ppm toluene. These concentrations bracket the threshold limit value of 100 ppm recommended by the American Conference of Governmental Industrial Hygienists (ACGIH).

Methods

SUBJECT SELECTION AND TRAINING

Forty two healthy students at the University of Michigan (aged 18–35), English speaking and with at least one year of college were studied. Candidates were excluded if there was a history of exposure to solvents or if they drank more than 18 g of alcohol a day (12·4 g is equivalent to 12 ounces of a 3·5% wt/vol beer).

Evaluation of performance was predicted on the assumption that the subject was highly motivated. To ensure this, a small payment ($35.00/day) was given. The odour of the toluene was masked with menthol crystal and no alcohol or caffeine was permitted the night before testing or during the three days of testing. A trial run on the battery was completed between 0700 and 0800 on the first day. It included manual and psychomotor response tests and omitted cognitive and memory tests, since cognitive tests were done only once and memory tests did not improve with practice in a pilot study. The tests were administered in a different order from later sessions in the presence of an instructor providing immediate feedback.

EXPOSURE CHAMBER

Exposures took place in two 6 × 6 × 7 ft chambers. The exhaust was set to 60 M^3/min. The temperature was kept constant at 25°C but humidity in the chamber did vary over the day.

EXPERIMENTAL DESIGN

The experimental design was a balanced two way scheme (exposure by day) for three permutations of one order of exposure in a standard 3 × 3 Latin square. Fourteen subjects were randomly assigned to each group. They were tested at different concentrations on each of the three days. The independent factors were (*a*) solvent concentration; (*b*) learning, defined as improvement in performance scores over the three days; and (*c*) the sequence of solvent exposure. The 0800 morning testing was repeated at 1145 and at 1600, representing three hour and seven hour exposures to toluene. Each subject's morning scores served as controls for the two sessions later in the same day. The control information was incorporated into the analysis by reporting the difference between a control and exposed score for each day, (am–pm). The difference (am–pm) was chosen since it is similar to testing before and after an eight hour workday. The study involved 126 exposures in 43 experimental days over three months.

Four subjects were tested a day in two behavioural chambers equipped with an intercom. The complete battery was given in the morning and afternoon. Four of the tests were also given at 1200. These tests were chosen for more frequent administration because they measured visuomotor response and manual dexterity, two aspects of performance adversely affected in previous solvent studies.[14] A fatigue checklist was given in the morning and afternoon.

At the end of the day symptoms and complaints were reported by subjects in response to the question "What are your reactions over the past four hours?" Observers also noted when subjects slept or dozed in the chamber in the morning and afternoon.

PERFORMANCE TESTS

A microcomputer (IBM PC) with a joystick (Ora Electronics) was used to administer the performance tests. The set of behavioural tests are modifications of two computer administered test batteries developed at the University of Michigan and Harvard University.[16 17 19] Contributions to the neurobehavioural evaluation system (NES) from our laboratory include a computerised version of the one hole test,[20] a short term memory digit span test,[21] a critical tracking task,[22] and a fatigue checklist.[23] The test battery covers seven human performance functions: verbal ability (vocabulary, Mill-Hill synonym test), verbal memory (Sternberg test, digit span), visual memory (pattern memory and Benton visual memory), perception (pattern recognition), visuomotor response (simple reaction time (SRT), a dynamic continuous performance test (CPT), a critical tracking test (CTT), and the symbol-digit test), manual dexterity (hand-eye coordination, the one hole test, finger tapping), mood (profile of mood scales (POMS), and a fatigue checklist). The test order is presented in table 1.

Table 1 *Order of administration of tests*

Tests	*Minutes*
One hole	10
Digit span	18
Mood	2
Pattern recognition	2
Pattern memory	2
Symbol digit	4
Hand-eye	3
Sternberg memory scanning	7
Simple reaction time	4
Finger tapping	3
Continuous performance test	5
Critical tracking	5
Benton visual memory test	4
Fatigue checklist	1
Total of minutes	70

EXPOSURE AND ANALYSIS OF TOLUENE IN AIR AND BREATH

A toluene and menthol solution was metered by a dyne pump (Milton Ray Co) and vapourised in a heated rippled glass distillation column. Menthol was maintained at 0·078 ppm. A constant negative pressure maintained the concentration in the chambers monitored by a Miran IA infrared analyser (Foxboro). The concentrations were independently confirmed by gas chromatography (GC) using toluene gas standards (Scotty) on a Varian 4000 GC with a flame ionisation detector fitted with a gas valve.

Eight hour menthol samples were also collected on XAD 7 custom made collector tubes (O.D. 8 cm, 1·10 cm, ID 6 cm, 200 mg in front, 100 mg in back, Supelco) using constant flow samplers (Dupont P2500A). Menthol was desorbed in 1:1 methanol and CS2 (reagent grade, Fischer), and analysed with a GC using a 2 metre column, 10% Carbowax 20 M on 100/120 mesh, WHP (Supelco).

Breath samples were collected by exhaling mixed air into a Teflon 12 l bag directly on leaving the chamber. Samples were collected before exposure at 0900 and after exposure starting at 1200 or 1700 at intervals of zero, five, ten, and 20 minutes. Samples were analysed by withdrawing air directly from the bag at 100 cc/min. Five millilitres were then trapped in the gas sampling loop, from which 2 ml were injected into the gas chromatograph. Each breath was analysed twice with a retention time (RT) of 4·1 minutes for toluene. The analytical method relied on comparing gas chromatograph results with a standard curve using certified gases at 10 and 50 ppm toluene (Matheson gases). The analysis was conducted within half an hour of collection.

STATISTICAL ANALYSIS

The effects of exposure, day of exposure (learning), and sequence of exposure were evaluated for the (am–pm) difference scores and on the scores at 1500 alone with an analysis of variance (ANOVA). The effects of toluene were similar for both dependent variables so that the increase in statistical power obtained by analysing the pm scores only did not alter the conclusions. Further, the variation in control data across three days was larger than the solvent effect; therefore, a within day difference score was the preferred dependent variable despite the added variation it introduced into the experiment.

The main hypothesis was that toluene caused a decline in human performance in a dose response manner. Differences between toluene concentrations were tested for consistency with a linear trend. Scheffe's 95% confidence intervals identified the specific significant differences between scores at 0 and 75 ppm or 0 and 150 ppm toluene.

A second analysis examined the linear association between breath concentration at 1700 and the (am–pm) performance scores. Each subject's performance score over the three exposures was regressed against their own breath values. An identical regression analysis was performed using the exposure categories 0, 75, and 150 ppm. A *t* test determined if the average slope across 42 subjects for both regressions differed

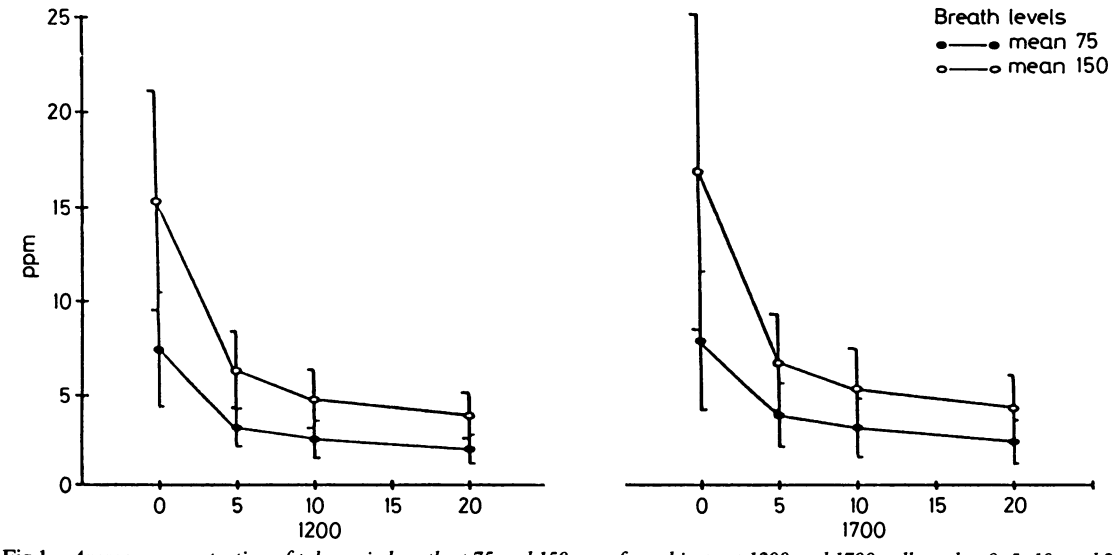

Fig 1 *Average concentration of toluene in breath at 75 and 150 ppm for subjects at 1200 and 1700 collected at 0, 5, 10, and 20 minutes after exposure.*

from zero. A separate t test also determined whether there was a difference between slopes, using each individual's slope divided by their own square root of the mean square error. To determine which slope was a better predictor of performance, the mean correlation coefficients of determination for respective regressions were compared for improvement in goodness of fit.

Symptoms were tallied by day and exposure for four categories: headache, irritation, tiredness, and anxiousness or frustration. An analysis of variance was performed on observed sleep patterns. Correlations between symptoms, mood, sleep, fatigue, age, sex, education level, and the performance score were carried out to rule out possibly confounded results.

Results

STUDY POPULATION

The small variability in age and education suggests that the population was homogenous (table 2). There were 21 men and women but no sex related effects were found. Since subjects correctly answered most of the questions on the vocabulary and Mill Hill synonym tests, the test scores were not important covariates in the ANOVA models.

EXPOSURE AND BREATH DATA

The average eight hour concentration of toluene in the chamber was either 74 ppm (SD = 3·7 ppm) or 151 ppm (SD = 7·5 ppm) (fig 1). Toluene breath levels appeared to be a linear function of toluene exposure. Breath levels at 75 ppm and 150 ppm toluene in air averaged 7·65 ppm (SD = 3·37 ppm) and 16·05 ppm (SD = 7·16 ppm). Twenty minutes after exposure, the average breath levels for both exposures dropped to 2·2 ppm (SD = 1·04 ppm) and 4·05 ppm (SD = 1·57 ppm).

RELIABILITY ESTIMATES

The test-retest reliability coefficients for control data (table 3) are the means across three morning test sessions and am v pm for each performance measure. The correlations within a day are usually greater than those across days, which is one reason for using the difference score within a day as the dependent variable.

Table 2 *Demographic features of the study population*

Age	Mean	22·5 (3·0)	
College	Years	3·5 (1·0)	
Dominant hand	Left	5	11·9%
	Right	37	88·1%
		Mean % correct (SD)	
Verbal skill	Vocabulary	86·8 (7·8)	
	Mill Hill synonym	86·4 (10·2)	

VERBAL SHORT TERM MEMORY

The Sternberg test was not affected by exposure or learning across three days, nor was there any interaction or effect in the order of exposure (table 4). The regression for the "yes" and "no" conditions were [yes msecs = 432·0 + 51·0 (item)] and [no msecs = 488·0 + 50·0 (item)].

The digit span scores were affected by toluene consistent with a linear trend (table 5, fig 2a). Scheffe 95% CIs confirmed a difference between 0 and 150 ppm but not between 0 and 75 ppm. At 150 ppm toluene the decrease in digit span was 0·44 (±0·99) digits or 6% based on an average control digit span of 7·56 (± 1·27). There were no solvent order, learning, or interaction effects.

Table 3 *Coefficients of determination for the average of three am and am v pm control scores within each control day*

	Reliability coefficients	
Performance measure	*am*	*am v pm*
Verbal short term memory		
Sternberg test:		
No intercept	0·5987	0·5140
No slope	0·4492	0·4888
Digit span:	0·6508	0·3932
Visual short term memory		
Benton visual memory:		
No correct		
Pattern memory:		
No correct	0·4532	0·6376
Latency	0·4243	0·8497
Perception		
Pattern recognition:		
No correct	0·5532	0·6466
Latency	0·4243	0·8497
Psychomotor functions		
Simple reaction time:	0·5588	0·6358
Continuous performance:	0·7128	0·8776
Symbol digit:		
No incorrect	0·0455	0·0908
Latency	0·0269	0·6976
Hand-eye coordination:		
RMSQ	0·5813	0·5538
Finger tapping:		
Right hand	0·5490	0·6605
Left hand	0·5309	0·6426
Alternating hand	0·3978	0·6214
Critical tracking test:	0·3906	0·5707
Manual dexterity		
One hole test:		
Position	0·7522	0·4885
Grasp	0·1290	0·5734
Move	0·4908	0·7989
Pin number	0·5543	0·6416
Reach	0·5662	0·8616
Mood scales		
Mood:		
Fatigue	0·5525	0·7416
Anger	0·6018	0·9213
Depression	0·6018	0·8406
Confusion	0·8121	0·8179
Tension	0·6949	0·8983
Fatigue checklist:	0·0948	0·1763

Table 4 *The Sternberg test, mood scales, and Benton visual memory test unaffected by toluene. Performance changes within a day (am–pm) with ANOVA p values for effects and trends for the day of exposure and ANOVA p values for effects and trends for concentration of exposure*

	Mean (am–pm) score (ppm)			*ANOVA p values*			
Mean over 3 days (SD)	*0*	*75*	*150*	*Day*	*Trend*	*Exp*	*Trend*
Verbal short term memory (Sternberg test)							
Yes Intercept							
20·72 (83·11)	16·0 (84·0)	16·0 (86·0)	29·0 (98·0)	0·518	0·296	0·747	0·533
Yes Slope							
4·54 (24·90)	5·2 (25·0)	5·0 (26·0)	3·5 (29·0)	0·474	0·219	0·951	0·769
Slope = msec/item	(+) sign intercept is an improvement						
Intercept = msec	(+) sign slope is a decrement						
Mood (profile on mood scales (POMS)							
Tension							
−0·01 (1·00)	0·00	0·01	−0·04	0·627	0·587	0·932	0·773
Depression							
0·04 (0·91)	0·03	−0·05	0·15	0·847	0·632	0·353	0·309
Anger							
−0·01 (0·99)	−0·04	−0·05	0·05	0·400	0·237	0·717	0·462
Fatigue							
0·10 (0·39)	0·20	0·05	0·05	0·718	0·838	0·444	0·315
Confusion							
0·16 (0·90)	0·24	0·11	0·00	0·625	0·576	0·552	0·390
Fatigue checklist							
0·88 (5·25)	1·66	0·11	0·00	0·625	0·576	0·552	0·390

(−) Sign mood scores are a decrement.

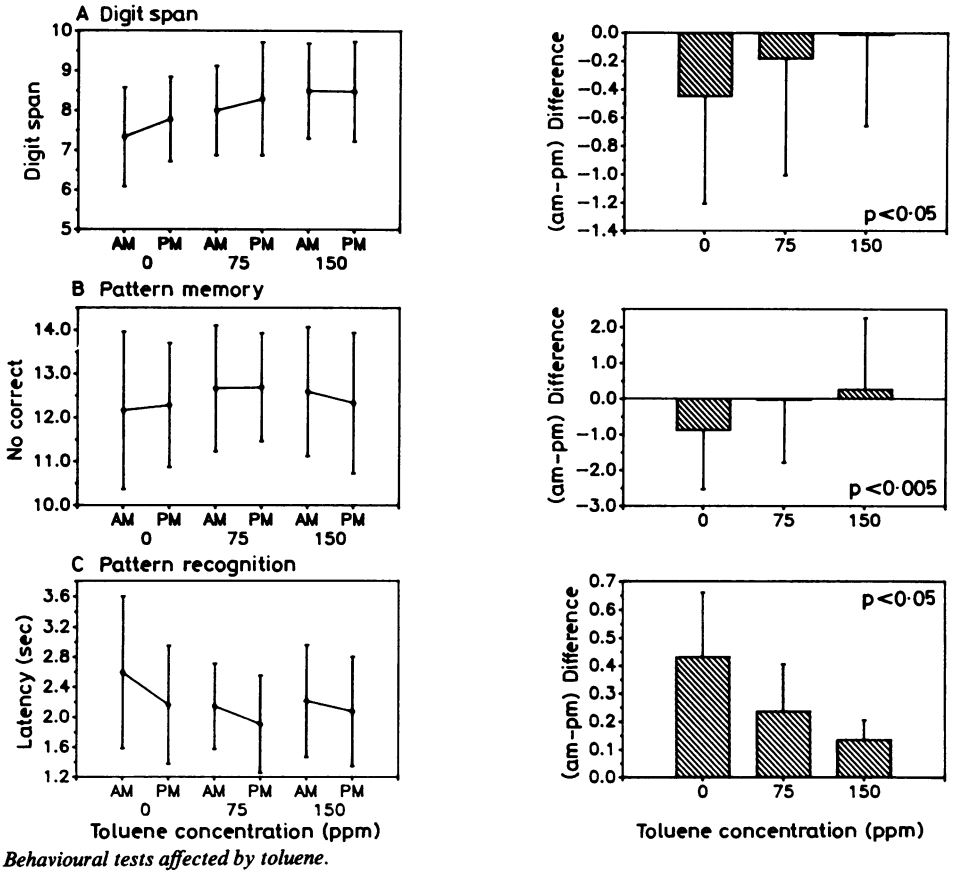

Fig 2 *Behavioural tests affected by toluene.*

Table 5 *Performance scores affected by toluene. Performance changes within a day (am-pm) with ANOVA p values for effects and trends for the day of exposure and ANOVA p values for effects and trends for concentration of exposure*

	Mean (am–pm) score (ppm)			*ANOVA p values*			
Mean over 3 days (SD)	*0*	*75*	*150*	*Day*	*Trend*	*Exp*	*Trend*
Verbal short term memory (digit span test)							
Digit span							
−0·21	−0·45	−0·18	−0·01	0·314	0·246	0·020	0·007
(0·74)	(0·76)	(0·82)	(0·68)				
Visual short term memory (pattern memory test)							
No correct							
−0·21	−0·88	−0·02	0·27	0·955	0·796	0·016	0·005
(1·78)	(1·74)	(1·76)	(1·81)				
Latency (secs)							
0·56	0·36	0·79	0·52	0·409	0·396	0·470	0·615
(1·12)	(1·60)	(1·95)	(1·88)				
Visual perception (pattern recognition test)							
No correct							
−0·19	−0·24	−0·24	−0·07	0·128	0·071	0·400	0·248
(0·68)	(0·74)	(0·64)	(0·71)				
Latency (secs)							
0·27	0·43	0·24	0·13	0·055	0·006	0·045	0·021
(0·47)	(0·54)	(0·38)	(0·63)				
Psychomotor ability (critical tracking test)							
Score							
−0·012	−0·033	−0·006	0·003	0·016	0·009	0·069	0·046
(0·067)	(0·105)	(0·059)	(0·059)				
Manual dexterity (one hole test)							
Pin No							
−1·4	−2·6	−1·7	−0·1	0·102	0·383	0·084	0·021
(4·9)	(4·3)	(6·3)	(5·9)				
Move (msecs)							
15·5	33·1	19·0	−5·6	0·406	0·991	0·003	0·001
(42·6)	(37·0)	(58·0)	(70·0)				

(+) Sign number is an improvement.
(−) Sign elapsed time is a decrement.

VISUAL MEMORY

The Benton visual memory test was not suitable for a repeated measures design since the stimulus was easy to memorise. Therefore the test was administered only at 0 and 150 ppm. The results of a paired t test performed on the number correct was not significant ($t = 1·7$, df = 41, $p < 0·1$). The pooled number correct was 10·96 (±1·01) or 91% correct (table 4).

The Pattern memory number correct was affected by toluene consistent with a linear trend (table 4, fig 2b). The Scheffe 95% confidence interval confirmed a difference between 0 and 150 ppm. At 150 ppm the decrement for the number correct was 1·14 (±2·43) or 9·4% compared with the average control score of 12·5 (±1·03).

VISUAL PERCEPTION

The pattern recognition latency for the number correct was affected by toluene consistent with a linear trend (table 5 and fig 2c). The Scheffe 95% CIs confirmed that the control score differed significantly from the 75 and 150 ppm scores. At 150 ppm the increase in the latency was 0·30 (±0·86) secs or 12·1% compared with the average control score of 2·31 (±0·60) secs.

PSYCHOMOTOR FUNCTION

Simple reaction time was not affected by toluene for the average of both hands or the left or right hand separately (table 6). There was, however, learning for the average of both hands on days 1, 2, and 3. The average improvement across days was 18 (±22·0) msecs or 5·4% from a baseline value of 336 (±36·0) msecs.

The continuous performance test was not affected by toluene (table 6). The mean response was 447 (±35·0) msecs. The mean numbers of omission and commission errors were 1·5 (±1·4) and 1·4 (±1·6). There was significant but inconsistent learning across the three days. The am–pm score improved 17 msecs (±14·0) or 3·2% for a baseline value of 431 (±31·0) msecs. The associated errors of omission and commission were not significantly affected by toluene but, like the response score, the number of commission errors was reduced consistent with a non-linear trend.

The number incorrect for the symbol-digit test showed significant differences between exposure conditions but in an irregular direction (table 6). Small errors in the number incorrect exaggerated the per cent change from the control score. At 75 ppm the number

Table 6 *Performance changes within a day (am-pm) with ANOVA p values for effects and trends for the day of exposure and ANOVA p values for effects and trends for concentration of exposure*

	Mean (am–pm) score (ppm)			*ANOVA p values*			
Mean over 3 days (SD)	*0*	*75*	*150*	*Day*	*Trend*	*Exp*	*Trend*
Simple reaction time (msecs)							
Right hand							
−5·3	−0·3	−5·3	−10·1	0·046	0·012 nl	0·411	0·152
(33·9)	(45·6)	(31·3)	(30·0)				
Left hand							
−3·2	1·8	−3·8	−7·5	0·026	0·034 nl	0·424	0·209
(34·8)	(31·6)	(39·8)	(39·4)				
Continuous performance test (msecs)							
4·17	5·44	2·42	4·66	0·109	0·014 nl	0·880	0·899
(26·78)	(25·0)	(29·0)	(30·0)				
Errors of omission							
−0·07	−0·02	0·09	0·28	0·127	0·165	0·614	0·492
(1·77)							
Errors of commission							
−0·27	−0·16	−0·21	−0·45	0·028	0·027 nl	0·773	0·364
(1·82)							
Symbol digit matching task							
No incorrect							
−0·05	−0·19	0·33	−0·28	0·510	0·405	0·032	0·021 nl
(1·16)	(1·38)	(1·49)	(0·67)				
Latency (sec)							
0·29	0·56	0·05	−0·62	0·000	0·000	0·486	0·896
(0·65)	(2·26)	(3·21)	(2·55)				
A hand-eye test (RMSQ)							
0·047	−0·051	0·050	0·142	0·032	0·010	0·704	0·396
(0·959)	(1·12)	(1·21)	(0·74)				
Finger tapping							
Right hand taps							
−1·38	−1·14	−1·82	−1·20	0·762	0·646	0·112	0·958
(6·18)	(7·46)	(6·11)	(6·30)				
Left hand taps							
−0·63	0·27	−2·26	0·10	0·047	0·041 nl	0·075	0·065
(5·62)	(5·42)	(6·34)	(5·68)				
Alt hand taps							
−2·46	−3·40	0·04	−4·03	0·118	0·469	0·232	0·699
(9·68)	(11·61)	(6·67)	(11·30)				

nl = Non-linear.
(+) Sign is an improvement.
(−) Sign for elapsed time is a decrement.

incorrect improved by 0·52 (±1·83) or 95%, and at 150 ppm the number incorrect got worse by 0·10 (±1·41) or 26% when compared with the average control score of 0·42 (±0·48). The Scheffe 95% CIs for the number incorrect confirmed that the three scores differed from each other. Its latency and SD were not affected by toluene but the latency did improve across the three days.

The hand-eye coordination test was not affected by toluene (table 6). The average root mean square was 4·61 (±1·17). There was significant learning across the three days. The root mean square improved by 0·56 (±0·87) or 12·5% based on the score from the first session of 4·46.

The finger tapping scores were unaffected by toluene (table 6). The mean number of taps over 10 seconds for the right, left, and alternating hands were respectively 58·8 (±7·24), 66·10 (±8·02), and 58·26 (±9·1). The left hand improved across the three days by 4·6 taps in 10 seconds or 7·6% based on a mean control score of 60·4 taps.

A critical tracking test (am–pm) score was barely affected by toluene. Borderline effects due to exposure to toluene were found consistent with a linear trend (table 4). The decrement at 150 ppm was 0·037 ±0·115 rads/sec (lambda) or 9·22% based on a control score of 0·398 ±0·076.

MANUAL DEXTERITY

Of the five variables of the one hole test, only "pin number" and the element "move" showed significant toluene effects without improvement across the three days, interaction, or solvent order effects (table 4, fig 3a, b). The Scheffe 95% CIs confirmed that the control (am–pm) difference score differed from the 150 ppm toluene score in a linear manner. At 75 and 150 ppm the pin number dropped 0·9 pins (±6·9) and 2·5 (±6·9) pins, a 1·9% and 6·5% decrease from an

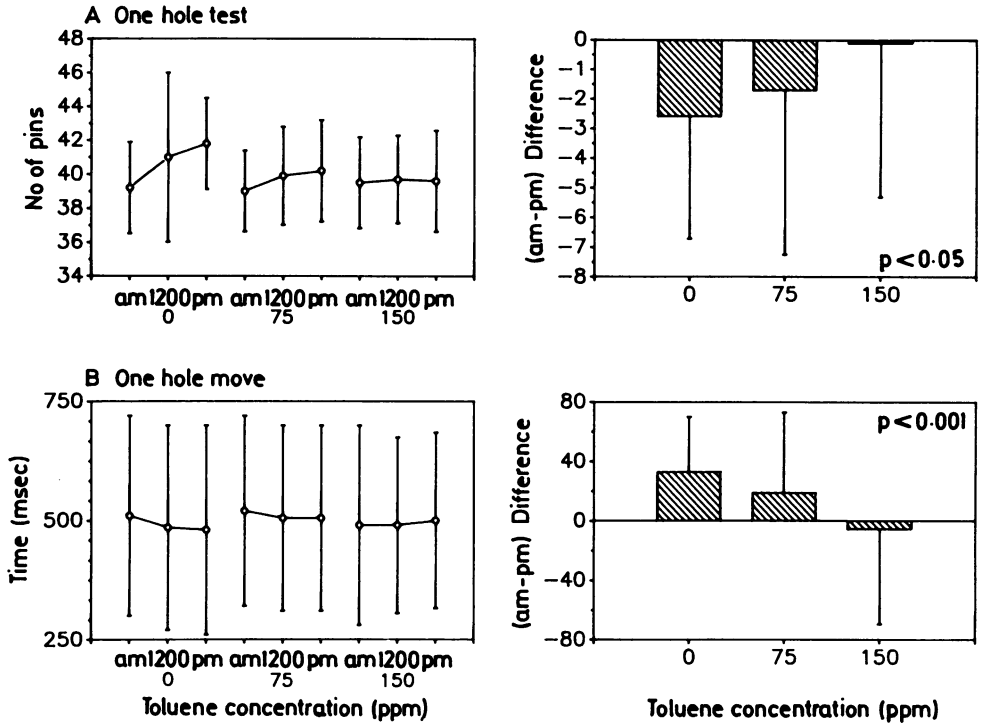

Fig 3 *One hole test affected by toluene.*

average baseline of 39 (±4·6 pins). For the element move, the time increased 14 (±62·0) msecs at 75 ppm and 39 (±73·0) msecs at 150 ppm, a 3·1% and 7·9% decrement from an average baseline of 508 (±59·0) msecs.

REGRESSION ANALYSIS

A comparison of the two slopes estimating dose, using breath levels or the categories 0, 75, and 150 ppm toluene, agreed with each other and were consistent with the ANOVA results (figs 4, 5). *T* tests performed on the standardised slopes never differed and the correlation coefficients of determination for both equations were similar, indicating little difference between the methods. There was one border line improvement in using breath with the critical tracking test score due to a decrease in residual error (fig 6). Otherwise, there was no advantage in using breath levels since the additional increment of variation explained by breath (less than 10%) was insufficient to increase the sensitivity of the tests.

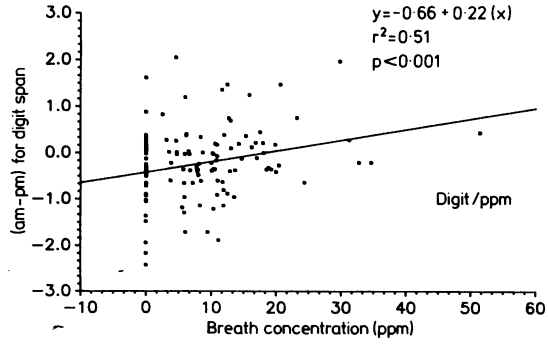

Fig 4 *Difference (am–pm) in digit span against toluene in breath.*

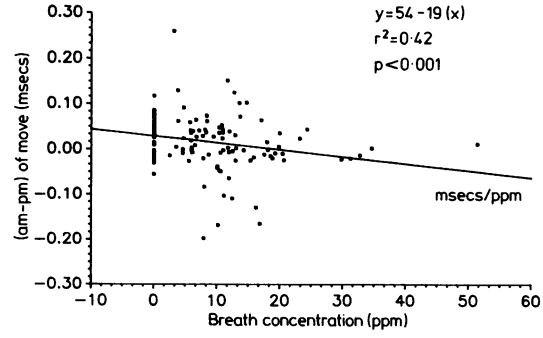

Fig 5 *Difference (am–pm) of one hole element "move" against toluene in breath.*

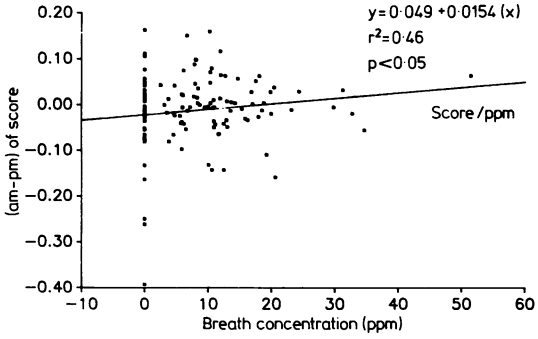

Fig 6 *Difference (am–pm) of a critical tracking score against toluene in breath.*

MOOD

The POMS scales and the fatigue checklist were not affected by toluene, learning, or by interaction (table 4). There was poor correlation between the fatigue checklist and the fatigue scale (−0·22). The R^2 never exceeded 0·2 between mood and performance test scores and therefore mood scales did not improve ANOVA models when used as covariates.

SUMMARY OF BEHAVIOURAL RESULTS

Table 7 gives a summary of the positive behavioural results found at 150 ppm for digit span, pattern recognition (latency), pattern memory (number correct), and manual dexterity. Most decrements cluster around 5–7%. A non-solvent related decrement of 26% was found for the latency of

Table 7 *Summary of the behavioural results*

	ANOVA		*Loss at 150 ppm*	*Control*
	Exposure	*Trend*	*(%)*	*(SD)*
Verbal short term memory				
Sternberg test:				
Digit span:	**	**	6·0	7·5 (1·27)
Visual short term memory				
Benton visual memory:				
Pattern memory:				
No correct	**	**	5·0	12·5 (1·03)
Latency				
Perception				
Pattern recognition:				
No correct				
Latency (secs)	**	**	12·1	14·63 (0·36)
Psychomotor functions				
Simple reaction time:				
Right hand				
Left hand				
Continuous performance:				
Symbol digit:				
No incorrect	*	*nl	26·0	0·42 (0·48)
Latency				
Hand-eye coordination (RMSQ):				
Finger tapping:				
Right hand				
Left hand				
Alternative hand				
Critical tracking (lambda):				
Manual dexterity				
One hole test:				
Pin No	*	**	6·5	39·2 (4·6)
Move (msec)	**	**	7·9	508·0 (59·0)
Mood scales				
Mood				
Fatigue checklist				

*p < 0·05; **p < 0·01.
nl = Non-linear.

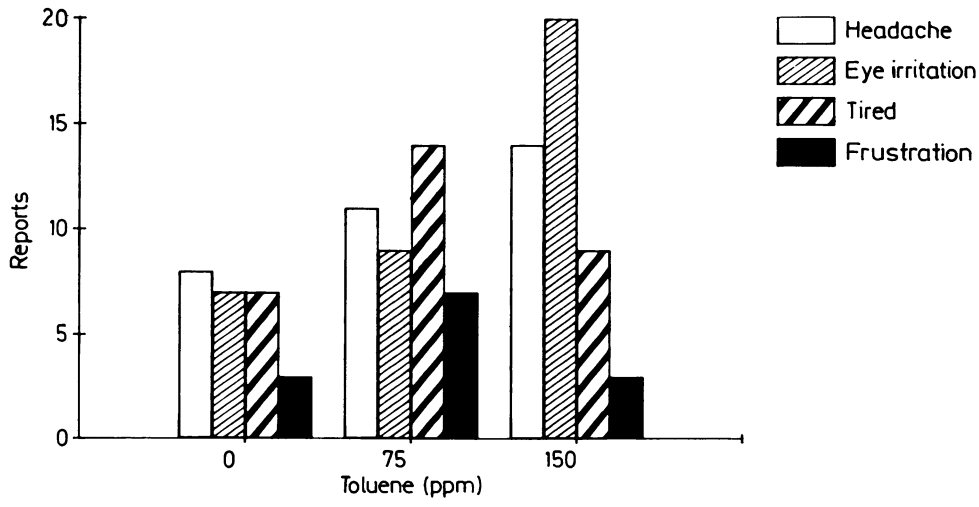

Fig 7 *Reports of symptoms at 0, 75, and 150 ppm toluene.*

encoding speed measured by the symbol-digit substitution test.

FACTOR ANALYSIS

A factor analysis was completed on all 42 (am–pm) performance variables and on a smaller set of 13 major response measures. A separate analysis was performed for each exposure condition at 0, 75, and 150 ppm toluene. The sensitivity of the larger analysis ranged from an eigen value of 0 to 2 permitting a maximum and minimum number of factors of 42 and 16. On the smaller subset of 13 variables the minimum number of factors was 6. In neither case was there a significant reduction in the data.

SYMPTOMS, SLEEP, AND MOTIVATION

The number of symptoms increases with increasing toluene exposure for headache, mucosal irritation, thirst,[3 8 12] and sleep (fig 7a, b). The reports were greatest on the first day of the experiment. Significantly more subjects slept in the afternoon, increasing from nine subjects (7%) on a control day to 18 subjects (14%) at 75 ppm toluene and 28 subjects (22%) at 150 ppm. The results were consistent with a linear trend ($p < 0.001$), but sleep did not confound toluene affected behavioural scores when used as a covariate in ANOVA models.

Half the subjects correctly guessed their order of exposure, meaning that the ability to recognise differences in exposure conditions was only partially controlled by masking the difference between toluene concentrations. It had no pronounced effect on performance, however, since there was no difference in results between the successfully blinded subjects and the non-blinded subjects.

Discussion

TOLUENE IN BREATH

The concentration of toluene in breath was used to confirm that the body burden of toluene was in equilibrium with the desired air levels when the subjects were tested. Under steady state conditions breath levels can be reported as a fixed percentage of the air concentration. We found that breath concentrations of toluene were proportional to air concentrations in agreement with previous research,[1 11] but the percentage was lower than expected. The expired air percentages were 10·5% at 75 ppm toluene and 11·2% for 150 ppm toluene. Results from animal studies indicate that 18% of an absorbed dose is excreted in breath[24] but human studies are more variable, ranging between 7%,[12] 10%, and 20%.[3 11]

The percentages in the present study are placed at the lower end of the range of reported estimates. One reason may be that we used a mixed air sample instead of an alveolar air sample. This will dilute the concentration of toluene. Using 18% as an ideal fixed percentage, the mixed air sample would have been diluted 37%. This percentage is based on the difference between the *observed* value at 150 ppm toluene from the *expected* value divided by the *expected* value (27 ppm–16·0 ppm)/27 ppm.

COMPARISON OF TWO ESTIMATES OF DOSE AS PREDICTORS OF BEHAVIOUR

It was hypothesised that breath samples would increase the ability to detect small changes in performance. Breath concentration is a more accurate estimate of dose correcting for intersubject variation in absorbance attributable to differences in gender,[25] body fat,[26] and ventilation rates.[1 11] Since the average coefficient of variation (SD/mean) for breath samples was 0·44, it was expected that considerable variation in performance scores could be explained by the range in breath levels.

The lack of improvement is explained by examining the difference in variation about the means for breath levels and the means for behavioural difference scores. The coefficient of variation of 0·44 for breath samples was always smaller than the coefficients of variation computed on control data for each test score. The smallest was 1·1 for the one hole element move (33·1 msecs/36·5 msecs). The largest was a 68·3 for the hand-eye coordination test (0·478 RMSQ/−0·007 RMSQ). The difference in variation about the mean for performance scores versus breath is greater than ten fold. Therefore, even when the analysis used toluene concentration in breath, the net reduction in the amount of variation is less than the performance variation across 42 subjects itself and is not sufficient to show differences in sensitivity between the two estimates of dose.

Biological monitoring in cross sectional behavioural studies should confirm the advantage of either measure of dose. Breath measurements would be preferable when intersubject variation in performance is affected by uptake which increases the sensitivity of the analysis. In the present study the average coefficient of variation of the control data was smaller than the variation in breath. This suggests that breath sampling might improve results in field studies where difference scores are not used, or where there are variable levels of exposure, or when more precise alveolar breath samples are collected.

BEHAVIOURAL DATA

The two verbal skill tests were used to evaluate whether this population was similar to other populations. For example, vocabulary differed between painters (EL Baker, private communication) in an industrial study (16/25 correct) and these college students

(22/25 correct). Differences in verbal skill may also affect other performance tests that require reading. In the present study the vocabulary and the Mill Hill synonym test were actually too simple with all the scores clustered at the high end.

The test-retest coefficients of determination for most performance scores were generally lower than those reported in field studies. This was not unexpected since the population was homogenous in age and education.

The two verbal short term memory tests, the Sternberg memory scanning test and a digit span test, were not equally sensitive to the effects of toluene. The Sternberg test was more variable (yes slope = 50 ± 23·03 msecs) than the digit span test (7·78 ± 1·07) and was not able to detect the effect of toluene. Digit span was affected by toluene in a dose response manner. The magnitude of the acute effect was reasonable since a loss of 0·44 digits (6%) is comparable with the effect of aging over 30 years (JM Schumaker, unpublished data).

The two visual memory tests, the Benton visual memory test and the pattern memory test, were also not equally sensitive to the effects of toluene. The Benton test was not significantly affected by toluene and was too simple, resulting in few errors for the number correct (96% ± 1% correct). Greater difficulty would increase the error and may make the test more sensitive to increasing solvent concentrations. The pattern memory test was more difficult than the Benton test (83·3% ± 10% correct) which may be why it was affected in a dose response manner. Two recent toluene and mixed solvent field studies also reported significant differences in a visual memory reproduction test.[27] Solvent sensitivity for the number correct was also shown with 40% nitrous oxide (FC Mahaney *et al*, in preparation) and with lead.[16]

Visual perception measured by pattern recognition was a simpler test than pattern memory (97·53% ± 2·6% correct), where the latency for the number correct was affected in a dose response manner. This result requires confirmation in future studies, since this test has not been used in other solvent studies. Visual perception, however, has also been adversely affected at 100 ppm toluene using the latency of the Landolt's ring test.[13 14]

There were six psychomotor tests. The visuomotor tests measuring speed such as simple reaction time, finger tapping, and a forced choice continuous performance test (CPT) were not significantly affected by toluene. Previous toluene studies are consistent with these results.[5-7 12] The hand-eye coordination test was also not significantly affected. The critical tracking task achieved borderline significance in a dose response manner. The symbol-digit substitution test that measures coding speed is complex and requires more judgment. The number correct was significantly affected by toluene but in an irregular manner. The improvement at 75 ppm toluene may be attributable to random error, and the 26% decrement found at 150 ppm may exceed the performance threshold for this test. This finding may be important because the symbol-digit WAIS subtest is reliable and commonly used in epidemiological studies. It is sensitive to solvents[27-29] and other neurotoxins[16] (and FC Mahaney, in preparation).

Manual dexterity was measured by the one hole test, designed to improve on the Purdue peg board test which uses identical pins but is much less specific.[20] The test measures the four elements "grasp," "move," "position," and "reach." These elements are components of a fifth measure "pin number." The task is to grasp a pin from a 2 cm indented bin, move it to a 0·5 cm hole, position the pin over the hole, release it, and reach for a new pin as rapidly as possible over one minute. The elements grasp, position, reach, and their fumbles were more variable than pin number or the element move, which is probably why they remained unaffected by toluene. At 150 ppm toluene, the element move was slightly more sensitive to toluene than its summary measure pin number (7·8 *v* 6·5%). Both measures have toluene decrements comparable with digit span and pattern memory. This test has not been administered in solvent behavioural studies before and its predecessor, the Purdue peg board test, has had inconsistent results. Test scores for the non-preferred hand were significantly affected in a mixed solvent exposure study in painters but not in a chronic single solvent exposure study of toluene in printers by the same investigators.[19 30] Significant effects, however, were found at 100 ppm toluene in chronically exposed printers used in an acute chamber study.[13 14]

AFTERNOON PM SCORES

A similar analysis using pm performance scores did not alter the profile of the effects of toluene. Decrements in performance were again found for pattern recognition, pattern memory, the one hole test, and the critical tracking test, with a new addition for the non-dominant hand of simple reaction time. A decrement in performance was not found for digit span.

LEARNING EFFECTS

Seven out of 11 performance tests improved with practice across the nine sessions. Motor response was quicker for five psychomotor tests: simple reaction time, continuous performance, hand-eye coordination, finger tapping (left hand), and symbol-digit substitution (latency). A motor component was also affected for the latency in the pattern recognition test, and for the element reach in the one hole test. Of all the

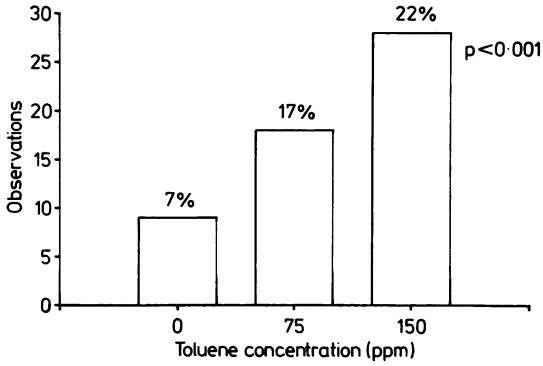

Fig 8 *Number of observations of sleep at 0, 75, and 150 ppm toluene.*

performance variables, only the latency for pattern recognition was simultaneously affected by exposure and learning without interaction between the two factors.

MOOD AND FATIGUE

The POMS scales and fatigue were initially measured as covariates in this study and were found to be independent measures since none of the scores was highly correlated with performance and since factor analysis did not reduce the number of independent measures accounting for the relation between performance scores.

SYMPTOMS AND SLEEP

Symptoms affected by toluene included headaches and eye irritation that increased in a dose response manner. Most reports occurred on the first day, indicating a higher level of awareness. Additional complaints included being tired, bored, or thirsty, consistent with previous studies.[4 6 14]

The most convincing evidence of toluene affecting the central nervous system was a dose response in the number of times subjects slept. The percentage increased from 7%, 14%, and 22% at 0, 75 ppm, and 150 ppm toluene (fig 8). The magnitude of effect indicates that objective behavioural tests may not be as sensitive to central nervous system effects as more general measures of feeling well.

STUDY DESIGN

One strength of the present study design was the ability to detect 5–10% changes in (am–pm) performance since performance variation exceeded the magnitude of the solvent effect. This degree of power was achieved by testing a large group of subjects where each subject was their own control. This study confirmed that manual dexterity and perception are affected by toluene (table 8) and presents new findings for visual memory and verbal short term memory. The effects are probably reversible at the biological threshold of response since the decrements cluster around 7% at 150 ppm toluene and do not usually exceed this level in a convincing manner.

Table 8 *Acute toluene chamber studies of non-industrially exposed subjects*

Author	*Country*	*Subjects*			*Exposure duration*	*Conc ppm*	*Behavioural tests and results*
		No	*Age*	*Sex*			
Dettingen *et al* 1942[4]	USA	3	35–53	M	8 h	600 300 100 50	Severe staggering and incapacitation Fatigue, muscular weakness, impaired coordination Moderate fatigue and sleepiness RR, PR, BP: ns
Ogate *et al* 1970[7]	Japan	5	23	M	7 h	200	Increased PR, SRT impaired BP: ns
Suzuki *et al* 1973[8]	Japan	5	18–22	M	6 h	200	Increased PR, RR, GSR, EEG, finger taps impaired Plethysmography; ns
Gamberale 1972[3]	Sweden	12	20–35	M	4(20')	700 500 300	CRT, perception impaired—identical no and spokes test Simple and choice RT impaired; perception tests: ns Simple RT impaired
Stewart *et al* 1973[11]	USA	14		M F	7½ h	100	Dual task visual vigilance and tone detection impaired Psychomotor tests: ns
Winneke *et al* 1982[10]	Sweden	18		M	3½ h	100	Critical flicker fusion, bisensory vigilance: ns Psychomotor tests: ns
Anderson *et al* 1982[14]	Sweden	16		M	6 h	0·40 100	Higher cortical function, visual perception, vigilance: ns visuomotor, manual dexterity, industrial simulations: ns
Anderson *et al* 1982[14]	Sweden	42e 42c		M M	6 h	100 0	Impaired manual dexterity and visual perception measure by Purdue peg board and Landolt's ring test
Dick *et al* 1984[12]	USA	18–30	18–38	M F	4 h	100	Impaired visual-vigilance Choice RT and pattern recognition: ns
Olson *et al* 1985[5]	Sweden	16	31	M	4 h	0·80 ± Xylene	Changes in mood and symptoms were reported vigilance, memory reproduction, simple and choice RT: n
Iregren *et al* 1985[6]	Sweden	12	22–44	M	4 h	0·80 ± Ethanol	Changes in mood and symptoms were reported Vigilance, memory reproduction, simple and choice RT: n

The fact that positive results were found on more stable measures and were more frequent in complex tasks suggests that performance decrements are more easily detected on tests in which there is little interday variation and that the sensitivity of the test battery depends in part on the specific study population. A group with more variability in performance or one that finds the tests simpler to perform would have to have larger decrements for them to be detected.

The ACGIH has adopted a lower threshold limit value (TLV) than the occupational safety and health administration's permissible exposure limit (PEL) of 200 ppm toluene. This study supports a lowering of the PEL because acute subjective and objective effects have been found at 75 and 150 ppm, bracketing the TLV of 100 ppm.

References

1 Astrand I, Ehrner-Samuel H, Kilbom A, Ovrum P. Toluene exposure I. Concentration in alveolar air and blood at rest and during exercise. *Work Environ Health* 1972;**9**:119–30.

2 Baker E, Fine L. Solvent neurotoxicity: the current evidence. *J Occup Med* 1986;**28**:126–8.

3 Gamberale F, Hultengren M. Toluene exposure II. Psychological functions. *Work Environ Health* 1972;**9**:131–9.

4 Von Oettingen W, Neal P, Donahue D, *et al.* The toxicity and potential dangers of toluene, with special reference to its maximal permissable concentration. *United States Public Health Services Bulletin* 1942;**279**:1–50.

5 Olson AB, Gamberale F, Iregren A. Coexposure to toluene and p-xylene in man: central nervous functions. *Br J Ind Med* 1985; **42**:117–22.

6 Iregren A, Akerstedt T, Olson BA, Gamberale F. Experimental exposures to toluene in combination with ethanol intake. *Scand J Work Environ Health* 1986;**12**:128–36.

7 Ogata M, Tomokuni K, Takatsuka Y. Urinary excretion of hippuric acid and m- or p-methylhippuric acid in the urine of persons exposed to vapours of toluene and m- or p-xylene as a test of exposure. *Br J Ind Med* 1970;**27**:43–50.

8 Suzuki H. Autonomic nervous responses to experimental toluene exposures in humans. *Sangyo Igaku* 1973;**15**:379–84.

9 Wellford TA. *Fundamentals of skill.* London: Methuen, 1968.

10 Winneke G. The behavioral effects of exposure to some organic solvents: psychological aspects. In: Junten J, ed. *Occupational neurology. Acta Neurol Scand* 1982;**66**(suppl 92):117–29.

11 Stewart RD, Hake CL, Forster HV, Lebrun AJ, Peterson JE, Wu A. *Toluene: development of a biological standard for the industrial worker by breath analysis.* Cincinnati: NIOSH, 1975. (DHEW-NIOSH contract report No 99-72-84.)

12 Dick RB, Setzer JV, Wait R, *et al.* Effects of acute exposure of toluene and methyl ethyl ketone on psychomotor performance. *Int Arch Occup Environ Health* 1984;**54**:91–109.

13 Anderson I, Lundqvist GR, Molhave L, *et al.* Human response to controlled levels of toluene in six-hour exposures. *Scand J Work Environ Health* 1983;**9**:405–18.

14 Anderson I. Exposure chamber studies. In: Cherry N, Waldron HA, eds. *The neuropsychological effects of solvent exposure.* (Proceedings of a symposium held at the London School of Hygiene and Tropical Medicine, 1982.) Havant, Hants: Colt Foundation, 1983:100–14.

15 Baker EL, Letz RE, Fidler AT. A computer administered neurobehavioral evaluation system for occupational and environmental epidemiology: rationale, methodology and pilot study results. *J Occup Med* 1985;**27**:206–12.

16 Baker EL, Letz RE, Fidler AT, Shalat S, Plantamura DL, Lyndon ML. A computer-based neurobehavioral evaluation system for occupational and environmental epidemiology: methodology and validation studies. *Neurobehav Toxicol Teratol* 1985;**7**: 369–77.

17 Maizlish NA, Fine LF, Albers JW, Whitehead L, Langolf GD. A neurological evaluation of workers exposed to mixtures of organic solvents. *Br J Ind Med* 1987;**44**:14–25.

18 Fidler AT, Baker EL, Letz RE. Neurobehavioral effects of occupational exposure to organic solvents among construction painters. *Br J Ind Med* 1987;**44**:292–308.

19 Cherry N, Venables H, Waldron HA. Exposure research in Britain in the neuropsychological effects of solvents. In: Cherry N, Waldron HA, eds. *The proceedings of a symposium held at the London School of Hygiene and Medicine, 1974.* Havant, Hants: Colt Foundation, 1983:136–51.

20 Salvendy G. Selection of industrial operators: the one hole test. *Int J Prod Res* 1975;**13**:303–21.

21 Smith PJ, Langolf GD, Goldberg J. Effects of occupational exposure to mercury on short term memory. *Br J Ind Med* 1983; **40**:413–9.

22 Jex HR, McDonnell JD, Phatak AV. A critical tracking task for manual control research. *IEEE Transactions on Human Factors in Electronics* 1966;7:138–45.

23 Pearson RG, Byars GE. *The development and validation of a checklist for measuring subjective fatigue.* Randolf Field, Texas: School of Aviation Medicine, 1956. (USAF report No 56–115.)

24 Teisinger J, Srbova J. Elimination of benzoic acid with the urine and its relation to the maximum tolerable toluene concentration in the air. *Archives des Maladies Professionnelles de Médecine du Travail et de Sécurité Sociale* 1955;**16**:216–20.

25 Carlsson A. Exposure to toluene: uptake, distribution and elimination in man. *Scand J Work Environ Health* 1982;**8**:43–55.

26 Nomiyama K, Nomiyama H. Respiratory retention, uptake and excretion of organic solvents in man: benzene, toluene, n-hexane, trichloroethylene, acetone, ethyl acetate, and ethyl alcohol. *Int Arch Arbeitsmed* 1974;**32**:75–83.

27 Irigren A. Effects on psychological test performance of workers exposed to a single solvent (toluene)—a comparison with effects of exposure to a mixture of organic solvents. *Neurobehav Toxicol Teratol* 1982;**4**:695–701.

28 Lindstrom K. Psychological performances of workers exposed to various solvents. *Work Environ Health* 1978;**4**:19–45.

29 Stokholm J, Cohr K. Acute subjective and objective reactions after 7 hours exposure to mineral turpentine. In: *Proceedings of the 26th Nordic Occupational Hygiene meeting.* Helsinki: Institute of Occupational Health, 1977:35–6.

30 Cherry N, Venables H, Waldron HA. The acute behavioural effects of solvent exposure. *J Soc Occup Med* 1983;**33**:13–8.