Journal of the International Neuropsychological Society (2005), **11**, 281–289.

DOI: 10.1017/S1355617705050332

An evaluation of persisting cognitive effects after withdrawal from long-term benzodiazepine use

MELINDA J. BARKER, KENNETH M. GREENWOOD, MARTIN JACKSON, AND SIMON F. CROWE

School of Psychological Science, La Trobe University, Victoria 3086, Australia

(RECEIVED January 14, 2004; 1ST REVISION July 22, 2004; ACCEPTED December 21, 2004)

Abstract

Twenty participants with self-reported long-term benzodiazepine use (mean 108 months) who had previously withdrawn from medication (mean 42 months) were administered a battery of neuropsychological tests. Each long-term user was case matched for age, sex, and education to two control participants who reported never taking benzodiazepines (those with and those without anxiety). The results indicated that long-term benzodiazepine use may lead to impairments in the areas of verbal memory, motor control/performance, and nonverbal memory but not visuospatial skills and attention/concentration. The length of abstinence (> 6 months) indicates that these impairments persist well beyond cessation of benzodiazepine use. However, observed impairments in the area of nonverbal memory were not solely attributable to benzodiazepine use and may be influenced by the elevated anxiety levels present in both the case and the anxious control group. (*JINS*, 2005, *11*, 281–289.)

Keywords: Benzodiazepines, Long-term effects, Neuropsychological assessment, Cognitive ability

INTRODUCTION

Despite questions regarding the safety of long-term benzodiazepine therapy, benzodiazepines remain among the most widely prescribed psychotropic medications worldwide (Australian Institute of Health and Welfare, 2000; Balter et al., 1984; Mellinger et al., 1984). Researchers in the Netherlands recently reported a prevalence of long-term use (more than 1 year) of .6% in a study of over 80,000 general practice patients (Zandstra et al., 2002). Concern is steadily increasing regarding the potential of benzodiazepines for dependency, significant withdrawal effects, and possible cognitive deficits (Ashton, 1986, 1995; Chen, 1990; Curran, 1986). The existing literature in the area is difficult to evaluate due to vastly conflicting results, study designs, and patient groups. However, previous research has suggested that long-term use of benzodiazepines is related to impairments in memory (Birzele, 1992; Curran, 1992; Curran et al., 1994; Massin-Krauss et al., 2002; Mintzer et al., 2001; Tata et al., 1994), attention and concentration (Birzele, 1992; Golombok, 1989; Golombok et al., 1988; Petursson et al., 1983), visuospatial skills (Bergman et al., 1980; Golombok et al., 1988; Sakol & Power, 1988; Tata et al., 1994), and numerous other cognitive functions (Aranko et al., 1983; Bergman et al., 1980, 1989; Birzele, 1992; Brosan et al., 1986; Gorenstein et al., 1994; Lucki et al., 1986; Petursson et al., 1983; Sakol & Power, 1988; Tata et al., 1994; Toenne et al., 1995). [See (Barker et al., 2003) for a recent review.] Some authors have also demonstrated a correlational relationship between benzodiazepine dose and observed deficits (Golombok et al., 1988; Tata et al., 1994).

Some researchers support the notion that the observed cognitive impairments improve following discontinuation (Salzman et al., 1992; Toenne et al., 1995); however, equally as many views to the contrary have been proposed suggesting that these patients display permanent or ongoing cognitive complaints (Gorenstein et al., 1994; Petursson et al., 1983; Tata et al., 1994). A recent meta-analytic evaluation of the literature indicated that previous long-term benzodiazepine users were impaired in all of the twelve cognitive areas assessed, as compared to controls (Barker et al., 2004a). Two further meta-analyses, focusing on improvement after discontinuation and impairment at long-term follow-up, suggest that, while some recovery of function was observed,

Reprint requests to: Professor Simon F. Crowe, School of Psychological Science, La Trobe University, Victoria 3086, Australia. E-mail: s.crowe@latrobe.edu.au

previous users display impairment in many areas some years after discontinuation (Barker et al., 2004b).

Clearly, there exists a need to further investigate the nature of cognitive impairment after long-term benzodiazepine use. Because the feasibility of large-scale studies examining a wide variety of cognitive areas is questionable, we have previously argued the need for a number of smaller, well-controlled studies, which present their data in a manner amenable to future meta-analysis (Barker et al., 2004a). The integration of findings from numerous studies that thoroughly investigate a small number of cognitive areas, include appropriate comparison groups, control for the effects of anxiety, and exclude patients with high alcohol or other drug use, seems a more practicable method for adding to our understanding of the nature of effects of long-term benzodiazepine use.

The present study attempts to address some of these issues by examining a small number of areas of cognitive functioning in previous long-term benzodiazepine users who have withdrawn and remained abstinent for at least 6 months, and were not taking other psychotropic medication. The results of these patients were then compared to two well-matched comparison groups, controlling for elevated anxiety levels.

METHOD

Participants

Participants (BZD group) were 20 people (7 males, 13 females), who reported having taken benzodiazepine medication regularly for a period of at least 12 months and had completely withdrawn from benzodiazepine medication at least 6 months prior to assessment. General exclusion criteria were heavy alcohol use (more than 15 standard drinks per week), illicit drug use (any period of regular use on self-report), head injury, stroke, other significant psychopathology, or current antidepressant medication use. Participants were recruited from an organization called Tranquilliser Recovery and New Existence (TRANX). One hundred and fifty past clients of TRANX were sent a letter outlining the research and asking for volunteers. Because the withdrawal status of the clients was unclear, the eligibility of these 150 was not known. Therefore, the clients were asked to contact the researcher if they wanted to participate, had been abstinent from benzodiazepines for at least 6 months, and were not currently taking antidepressant medication. Approximately 6 weeks after the initial letter, reminder letters were sent. A total of 32 letters were returned "not at this address". Of the 33 past TRANX clients who contacted the researcher (a response rate of 28%), nine were still taking benzodiazepine medication, five were taking antidepressant medication, two had been regular narcotic users, one had suffered a significant head injury, and one was ineligible due to high alcohol intake. The remaining 15 were eligible and were recruited into the study. In addition, 16 people contacted the researcher after an article about the research was published in a statewide newspaper. Five of these were eligible for the study and were recruited. The remaining eleven were ineligible due to current benzodiazepine use (3), current antidepressant use (3), head injury (1), aneurysm (1), epilepsy (1), heavy alcohol use (1), or had not used benzodiazepine for longer than 1 year (1).

Each of the 20 participants was matched closely to two controls for age (within 5 years), sex, and education (within 2 years). The first group of controls (ANX group) comprised people who had been diagnosed with an anxiety disorder and reported never having taken benzodiazepine medication regularly for any period, and were not taking antidepressant medication. These participants were recruited from an anxiety support group *via* a quarterly newsletter (3 participants) and from an advertisement in the same statewide newspaper as the BZD group advertisement was placed (17 participants). The same general exclusion criteria applied. Of the 47 phone calls that were received following the newspaper article, thirty were ineligible for the following reasons: did not match a BZD group participant due to age (8), sex (1) or education (6), previous benzodiazepine use (3), current antidepressant use (9), current benzodiazepine use (2), or did not suffer from anxiety (1).

The second group of controls (NML group) was an incidental sample of people who had never been diagnosed with an anxiety disorder (self reported), had never taken benzodiazepine medication regularly (self reported), and were not taking antidepressant medication. The same general exclusion criteria applied. All participants were reimbursed $20 for their time and/or travel costs.

Test Battery

Five of the cognitive areas, identified in the follow-up meta-analysis (Barker et al., 2004b) as having a moderate or greater effect size, were chosen for further assessment. These were attention/concentration, motor control/performance, nonverbal memory, verbal memory, and visuospatial skills. A measure of anxiety and measures of general intelligence were also used. Each test was categorized as measuring a certain cognitive domain based upon test descriptions in two neuropsychological texts (Lezak, 1995; Spreen & Strauss, 1998). Individual tests in the test battery and the cognitive domain assessed are presented in Table 1.

The demographic information collected included date of birth, high-school educational level, post-high-school education (including trade certificates, diplomas, and tertiary degrees), occupation, weekly number of alcoholic drinks, medical history information (i.e., head injury, stroke, etc), and other medications used. The additional information collected on the BZD group included the condition for which the medication was prescribed, year started, prescription source, type of benzodiazepine, daily dosage, length of use, month last taken, and nominated reason for attending TRANX or wishing to discontinue.

Table 1. Test battery items and cognitive domains assessed

Cognitive domain	Test
Verbal memory	Wechsler Memory Scale III—Logical Memory I & II (Wechsler, 1997)
Motor control/performance	Purdue Pegboard (Tiffin, 1987)
Nonverbal memory	Visual Spatial Learning Test (VSLT) (Malec et al., 1992); Austin Maze (Walsh, 1985)
Visuospatial skills	Benton Judgement of Line Orientation—Form H (JLO) (Benton et al., 1983); WASI—Block Design subtest (Wechsler, 1999); Gestalt Closure Test—Kaufman ABC (Kaufman & Kaufman, 1983)
Attention/concentration	Trail Making Test–Part A (TMT-A)(Reitan & Wolfson, 1985)
Anxiety	State-Trait Anxiety Inventory (STAI) (Spielberger, 1983)
General intelligence	National Adult Reading Test–2nd Edition (NART–II) (Nelson & Willison, 1991); Wechsler Abbreviated Scale of Intelligence (WASI)—Full Scale IQ (Wechsler, 1999)

Testing Procedure

Tests from the test battery were administered to participants in randomized order by assigning each test a number and using a random number table. The exception to this was Logical Memory I and the Visual Spatial Learning Test (VSLT), which were administered within the first five tests to ensure the delayed recall aspects of these tasks did not unnecessarily prolong the session. All but two participants completed the entire battery in two, 1-hr sessions on the same day. On those two occasions, the assessment battery was administered over two 1-hr sessions on separate days.

Data Analysis

Where available, raw scores were converted to age-related transformed scores (i.e., scaled scores, percentiles or T-scores) from each test's normative data. For two tests (Gestalt Closure Test and VSLT) this data was not available and therefore the raw scores were used in the analysis. Scores of each of the three groups on each of the test battery items were compared using a single-factor, repeated-measures ANOVA. Due to potential problems with sphericity, the Huynh-Feldt correction was used. Given the number of items in the test battery, a Bonferroni-corrected p value of .002 was used in assessing significance. On those items where a significant difference between groups was detected, *post hoc* investigations were conducted using the Fisher-Hayter procedure (Kirk, 1995). Effect sizes were calculated according to the method described by Rosenthal (Rosenthal, 1991). A negative effect size indicates that the BZD group's performance was worse than the control group's performance.

RESULTS

Subject Characteristics

Patients and controls were well matched on age, sex, and education level. Group characteristics are presented in Table 2. Patients were matched to each control within 5 years of age, resulting in no significant difference between groups, $F(1.99, 37.80) = 1.03$, $p = .367$. There was also no significant difference in the occupational status of participants in each group, coded using the ANU3_2 Scale (McMillan & Jones, 2000), $F(2.0, 36.0) = 2.27$, $p = .118$. There was also no significant difference between groups on the number of standard alcoholic drinks consumed per week $F(1.95, 36.99) = 2.44$, $p = .102$. Differences between groups on the measures of general intelligence (WASI FSIQ) and pre-morbid IQ estimation (NART–II) just failed to reach the required significance level, $F(1.73, 32.79) = 7.77$, $p = .003$, and $F(1.76, 33.36) = 7.25$, $p = .003$ respectively.

Repeated-measures ANOVA revealed a significant difference between the groups on STAI scores for the Trait Anxiety scale only, $F(1.77, 33.69) = 13.67$, $p < .001$. *Post hoc* analyses using the Fisher-Hayter procedure indicated that the NML group achieved significantly lower Trait anxiety scores than both the BZD group, $q^H = 4.72$, $p < .01$, and the ANX group, $q^H = 7.29$, $p < .01$. There was no significant difference between the BZD and ANX groups $q^H = 2.58$, $p > .05$.

The conditions for which the BZD group were originally prescribed medication were generalized anxiety disorder (7), panic attack disorder (5), insomnia (2), chronic pain (2), depression (1), jetlag (1), posttraumatic stress disorder (1), and as a muscle relaxant (1). Fourteen of the BZD group participants were prescribed their medication by a general medical practitioner, five by a psychiatrist, and one by another type of medical professional. The most common type of benzodiazepine taken was diazepam (9), followed by alprazolam (3), clonazepam (3), oxazepam (2), flunitrazepam (1), clorazepate (1), and temazepam (1). The mean diazepam equivalent for the BZD group was 33.1 mg (*SD* 32.8, range 7.5–160). The mean length of use in months was 108.5 (*SD* 95.5, range 12–348) and mean length of abstinence in months was 42.2 (*SD* 50.8, range 6–174.5).

The reasons nominated for wanting to discontinue by the participants in the BZD group were health reasons or worried about side effects (12), it had stopped working (3), was required to withdraw to enter cognitive behavior therapy (2), felt more depressed (2), and didn't feel they were functioning well (1).

Table 2. Participant characteristics for each group

Characteristic	Group		
	NML	ANX	BZD
Mean age (*SD*)	49.2 (11.3)	48.9 (10.9)	49.8 (12.1)
Male/female ratio	7/13	7/13	7/13
Years of education			
7–8	1	1	1
9–10	3	3	3
11–12	16	16	16
Post secondary education			
None	6	7	6
Trade certificate	2	2	2
Diploma	6	6	6
Tertiary degree	6	5	6
Mean ANU3_2 score (*SD*)	44.5 (18.7)	35.5 (13.5)	40.5 (13.6)
Mean weekly alcohol intake, standard drinks (*SD*)	5.9 (4.1)	4.9 (5.6)	2.8 (4.8)
Mean WASI FSIQ score (*SD*)	113.75 (8.0)	106.7 (8.2)	105.1 (9.6)
Mean NART–II IQ estimation (*SD*)	108.9 (6.9)	113.4 (4.0)	106.8 (7.5)
Mean STAI percentile score (*SD*)			
State	39.1 (19.5)	66.4 (26.2)	68.7 (28.7)
Trait*	52.8 (22.5)	90.4 (13.0)	77.1 (27.8)

*$p < .00$.

Individual Test Results

Table 3 displays the means, standard deviations, and significance levels of individual test scores for each group, as well as the effects sizes for each between-group comparison.

Verbal memory

The NML group performed significantly better than the BZD group on Logical Memory I and II. Similarly, the ANX group's performance was significantly better than the BZD group on both measures. The effect sizes for these differences were all large. There was no significant difference between the ANX and NML groups.

Motor control/performance

Significant differences were found across all four measures of the Purdue Pegboard. Both the ANX and NML groups performed significantly better than the BZD group. There were no significant differences on any of these measures between the ANX and NML groups. All of these effect sizes were large in magnitude.

Nonverbal memory

Of the four measures of the VSLT (correct designs, correct positions, correct positions and designs, incorrect designs), the BZD group performed significantly worse than both the ANX group and the NML group on the number of correct positions measure. Both of these effect sizes were large in magnitude. There was no difference between the NML and ANX groups. In addition, there were no significant differences on any of the remaining three measures of the VSLT between any of the groups.

In contrast, on the Austin Maze, the NML group achieved a significantly higher percentile score than both the ANX group (moderate effect size) and the BZD group (large effect size). There was no difference between the BZD group and the ANX group.

Visuospatial skills

There was no significant difference observed between any of the groups on the Benton Judgement of Line Orientation (JLO), the WASI Block Design subtest, or the Gestalt Closure Test.

Attention/concentration

There was no significant difference observed between any of the groups on the Trail Making Test Part A.

Cognitive Category Results

The mean effect size difference between the BZD and the combined control groups was calculated for each cognitive category and are presented in Table 4. In addition, the mean effect sizes for each category assessed were compared to the corresponding category effect sizes found in the previous meta-analysis conducted at long-term follow up (Barker et al., 2004b). These data are presented in order of decreasing effect size to facilitate comparison of the order of largest to smallest effect size for the cognitive areas studied. As can be seen from Table 4, with the exception of nonverbal memory mov-

Table 3. Means and standard deviations of individual test scores for each group and effect sizes for between-group comparisons

Cognitive category	Mean (*SD*)						Effect Sizes *d* (95% CI; upper, lower)		
	NML group		ANX group		BZD group		NML *vs.* ANX	NML *vs.* BZD	ANX *vs.* BZD
Verbal Memory									
Logical Memory I—scaled score	12.15	(1.31)	11.35	(2.52)	8.60	(2.48)	−.40 (−1.02, .24)	−1.80** (−2.48, −1.03)	−1.10** (−1.74, −.41)
Logical Memory II—scaled score	12.60	(1.39)	11.85	(2.21)	9.25	(2.51)	−.40 (−1.02, .23)	−1.66** (−2.33, −.90)	−1.10** (−1.74, −.41)
Motor Control/Performance									
Purdue Pegboard—right hand—percentile	56.60	(16.55)	51.50	(23.35)	28.65	(23.41)	−.25 (−.86, .38)	−1.37** (−2.03, −.66)	−.98** (−1.61, −.30)
Purdue Pegboard—left hand—percentile	47.75	(20.23)	48.00	(20.61)	25.25	(18.46)	.01 (−.61, .63)	−1.16** (−1.81, −.47)	−1.16** (−1.81, −.47)
Purdue Pegboard—both hands—percentile	54.85	(10.21)	52.80	(13.69)	24.35	(19.92)	−.17 (−.79, .45)	−1.93** (−2.63, −1.14)	−1.66** (−2.35, −.92)
Purdue Pegboard—assembly trial—percentile	60.50	(15.60)	56.85	(22.47)	36.55	(25.54)	−.19 (−.81, .44)	−1.13** (−1.77, −.44)	−.84** (−1.47, −.18)
Non-Verbal Memory									
VSLT—correct designs—raw score	32.40	(2.14)	31.45	(2.11)	29.30	(3.64)	−.45 (−1.07, .19)	−1.04 (−1.68, −.36)	−.72 (−1.35, −.07)
VSLT—correct positions—raw score	28.80	(2.59)	26.95	(3.75)	21.25	(5.92)	−.57 (−1.19, .07)	−1.65** (−2.33, −.91)	−1.15** (−1.79, −.46)
VSLT—correct positions and designs—raw score	26.55	(3.33)	23.10	(4.24)	18.55	(6.53)	−.90 (−1.54, −.24)	−1.54 (−2.22, −.81)	−.83 (−1.45, −.17)
VSLT—number of incorrect designs—raw score[a]	1.50	(1.76)	2.40	(1.88)	2.50	(1.91)	−.49 (−1.11, .14)	−.54 (−1.16, .10)	−.05 (−1.74, −.41)
Austin Maze—total errors at ten trials—percentile	43.75	(13.75)	32.00	(20.99)	22.35	(16.72)	−.67* (−1.28, −.01)	−1.40** (−2.06, −.68)	−.05 (−.67, −.57)
Visuospatial Skills									
Benton JLO—Percentile	66.10	(13.53)	54.15	(24.00)	42.10	(28.67)	−.61 (−1.23, .03)	−1.07 (−1.71, −.39)	−.46 (−1.07, .18)
WASI—Block Design—T score	57.45	(5.79)	55.05	(7.09)	51.60	(8.98)	−.37 (−.99, .26)	−.78 (−1.40, −.12)	−.43 (−1.04, .21)
Gestalt Closure Test—raw score	21.55	(1.70)	20.00	(4.23)	19.35	(3.27)	−.48 (−1.10, .16)	−.84 (−1.47, −.18)	−.17 (−.79, .45)
Attention/Concentration									
Trail Making Test—Part A—time in seconds[a]	24.80	(5.15)	29.05	(5.96)	38.05	(21.22)	−.76 (−1.39, −.11)	−.86 (−1.49, −.19)	−.58 (−1.20, .07)

Note. Effect sizes of 0.2, 0.5, and 0.8 are considered small, medium and large in magnitude (Cohen, 1988).
[a]Sign of effect size reversed.
**Post Hoc* analysis using Fisher-Hayter procedure, $p < .05$.
***$p < .01$.

Table 4. Mean effect sizes for each cognitive category presented in decreasing effect size order for both the current study and the previous meta-analysis conducted at long-term follow-up

Meta-analysis conducted at LT follow-up		Current study*	
Category	Weighted effect size (*d*)	Category	Mean effect size (*d*)
Verbal memory	−1.50	Verbal memory	−1.43
Motor control/performance	−.62	Motor control/performance	−1.33
Visuospatial skills	−.49	Nonverbal memory	−.86
Attention/concentration	−.43	Attention/concentration	−.85
Nonverbal memory	−.26	Visuospatial skills	−.69
Overall	−.48	Overall	−.92

*Statistically significant differences were found between BZD and combined control groups on measures of verbal memory and motor control/performance.

ing from 5th to 3rd place, the order of effect size magnitude found is similar between the two investigations.

DISCUSSION

The results of this study indicate that previous long-term benzodiazepine users, assessed after at least 6 months of abstinence, continue to display cognitive deficits in a number of areas, compared to matched controls. Significant, moderate-to-large effect sizes were observed within the areas of verbal memory, motor control/performance, and nonverbal memory when comparing the BZD and NML groups. Significant differences were also found between comparison of the BZD and ANX groups on these measures.

The observation that the ANX and NML performed similarly on most measures and significantly better that the BZD groups implicates long-term benzodiazepine use as the most plausible explanation for the difference. Elevated anxiety levels could not account for poorer performances observed by the BZD group, as the both BZD and ANX groups displayed very similar State Anxiety levels. Gorenstein et al. (1995) also found their group of previous long-term benzodiazepine users, assessed at long-term discontinuation (average 10 months), to be impaired across a number of cognitive areas, including verbal memory, when compared to both an anxious control group and a normal control group. Furthermore, the past benzodiazepine user group and anxious control group did not differ significantly on anxiety scores, which were also significantly different from the normal control group. These authors argued that the absence of significant differences between the performance of anxious benzodiazepine-free controls and normal controls on all measures (except a manual dexterity task) indicates that poor performance in the past benzodiazepine user group could not be attributed to anxiety (Gorenstein et al., 1995). Similarly, other researchers have discounted anxiety as the likely cause of poor performance in long-term benzodiazepine users when the effects of anxiety are accounted for (Golombok et al., 1988).

On only one measure, the Austin Maze (nonverbal memory), did the ANX and NML groups differ. In addition, on this measure there was no difference between the BZD and the ANX groups, indicating that elevated anxiety levels may have impacted negatively on performance. Qualitatively this was certainly the case, with many of the participants in the both the BZD and the ANX groups reporting that this task, which includes a buzzer to indicate an incorrect move, was particularly anxiety provoking.

While the overall results of the current study are difficult to compare to previous studies due to the variety of study designs and cognitive areas examined, support for impairments in the areas assessed in this study is provided by the following studies: verbal memory (Tata et al., 1994), motor control/performance (Gorenstein et al., 1995; Petursson et al., 1983), nonverbal memory (Birzele, 1992), visuospatial skills (Tata et al., 1994; Golombok et al., 1988; Sakol & Power, 1988), and attention/concentration (Birzele, 1992; Golombok, 1989; Sakol & Power, 1988). The current study did not detect significant differences in performance in the areas of visuospatial skills and attention/concentration between the BZD group and controls. However, despite significance levels failing to reach the Bonferroni-corrected *p* value of .002, the observed trends indicated that the BZD group performed poorer than the controls in these areas, with significance levels consistently less than .05.

The findings of this study are supported by a comparison of the areas of impairment found to those in the previous meta-analysis (Barker et al., 2004b) by arranging the areas investigated in order of most to least impairment indicated by effect size magnitude (Table 4). The previous meta-analysis integrated the results from a number of studies that reassessed withdrawn long-term benzodiazepine users after at least 6 months of abstinence. By comparing the BZD group to the combined control group, calculating a mean

effect size for each cognitive area, and then ordering the effect sizes from greatest to smallest, a very similar pattern is evident. For example, large differences were consistently observed in performance between groups on verbal memory tasks in both the current study and the previous studies included in the meta-analysis, resulting in the largest effect size observed in the cognitive area of verbal memory.

The effect sizes observed in this study were larger in magnitude to those found in the previous meta-analyses for all categories except verbal memory. A possible explanation for this is the high level of control and the choice of dependent measures. The tests used in this study were chosen based on their sensitivity to detect deficits in particular areas, whereas the category results form the previous meta-analysis are based on results from a variety of measures. The similar effect size magnitude found for the category of verbal memory does not follow this argument but may either reflect a ceiling effect, or be due to the minimal variation of measures used to assess verbal memory. Further investigation revealed that, of the studies in the previous meta-analysis that investigated verbal memory, most used a very similar type of story memory task to each other and to the current study.

The results of this study, combined with the results from the previous meta-analyses indicate that, in some areas, differences do exist between the performance of previous long-term benzodiazepine users and those who have not used this medication. The failure of some previous studies to detect differences may again be attributable to the lower level of control and the small sample sizes often employed, resulting in the small effect sizes and nonsignificant differences reported. However, by combining all of the positive and negative effect sizes from the available previous research, the magnitude of the resulting category effect sizes support the existence of differences between previous long-term benzodiazepine users and controls (Barker et al., 2004a, 2004b) .

Tata et al. (1994) have previously pointed out the well-established link between acute benzodiazepine administration and adverse effects on memory and arousal, and attempted to link this to the high density of benzodiazepine receptors found in the hippocampus and reticular formation (Wolkowitz et al., 1987). The two largest differences observed in the current study were in the areas of verbal memory and motor control/performance—functions largely subserved by the hippocampus and reticular formation. However, further research is required before any conclusions may be drawn on the mechanism of persisting deficits, or the relationships between the pattern of deficits observed after withdrawal and physiological differences in the distribution of benzodiazepine receptors in the brain.

While this paper addresses some of the methodological criticisms of previous research in this area, a number of limitations remain. First, the small sample size of twenty in each group resulted in limited power to detect differences in some of the areas assessed. Furthermore, the Bonferroni correction applied, due to the number of analyses conducted, resulted in an extremely conservative test of significance. As a result, some real differences may have not been detected. The small sample size also precluded a comprehensive regression analysis of the contribution of length of use, dosage, and length of abstinence. This type of analysis is the logical and necessary next step in the body of research in this area.

The small sample sizes in this study are partially attributable to the method of recruitment and the strict exclusion criteria that were applied. In the process of attempting to address one of the previous criticisms—that some studies tend to include patients with high alcohol or current antidepressant use—the difficulty of finding previous long-term benzodiazepine users *not* currently taking other medication became apparent. It seems that a common practice for long-term benzodiazepine users who wish to withdraw is to replace their benzodiazepine medication with antidepressant medication. Many of those excluded from the study who had withdrawn, experienced a return of their initial symptoms and were prescribed a selective serotonin reuptake inhibitor (SSRI) to treat their anxiety. The increasingly popular SSRIs were often considered the lesser of two evils by both patients and their medical professionals.

Second, the recruitment method employed may have resulted in a biased sample. Most of the patients who attend TRANX do so to gain assistance in withdrawing from their medication. However, when mailing out to past TRANX clients, it was not known whether they had successfully withdrawn, their current withdrawal status, their current medication regime, or whether they had even been taking a benzodiazepine medication. This would have resulted in the invitation letters being sent to a number of people who were actually not eligible to participate. Of those who responded to the mailout, less than half were eligible for a variety of reasons. The second method of recruitment involved a newspaper advertisement. Consequently, the resulting group who were recruited into the study may have been a subgroup of previous users who were experiencing cognitive problems and were therefore more likely to respond to the letter or advertisement in order to undergo some investigation of their difficulties. The majority of the enrolled participants cited concerns about health or side effects as the predominant reason for wanting to discontinue. Furthermore, it may be the case that the patients who successfully withdraw from benzodiazepine medication do so because they are insightful of cognitive impairments, while those not experiencing any negative effects are less inclined to attempt to withdraw.

Third, it is not possible to comment on the preexisting cognitive functioning of the past benzodiazepine users. Lucki et al. (1986) have previously pointed out the difficulty that exists in estimating what a participant's performance may have been prior to, or without, long-term benzodiazepine use. The present study used an estimate of premorbid intelligence; however, premorbid performance on measures such as memory or psychomotor performance is impossible to determine. While the groups in the present study did not

differ statistically on the measures of premorbid intelligence (NART–II) and general intelligence (WASI FSIQ), these differences almost reached significance and should be acknowledged. Despite controlling for education, the NML group appeared to obtain a higher WASI FSIQ than both the BZD and ANX groups. In addition, the ANX group appeared to achieve a higher NART–II score than both the BZD and NML groups. These two measures would be expected to correlate highly and therefore this finding is considered highly unusual. One possible explanation is that the NART–II predicts WAIS–R FSIQ whereas the WASI is an estimate of WAIS–III FSIQ, with the older normative data tending to yield comparatively higher IQ scores.

Fourth, while differences between groups in the category attention/concentration did not reach significance, the use of a measure more sensitive to attentional dysfunction than the Trail Making Test–Part A would have been more appropriate, and may have revealed some differences.

Fifth, toxicology testing was not available during this study, and therefore the researchers relied on self-reported status of continued abstinence, and current use of illicit drugs or alcohol. While there was no way to confirm these states, there was no incentive to be other than truthful and the research was conducted completely independently of any connections the participant may have had to treatment facilities.

Finally, the same examiner conducted all of the psychometric testing, and due to the study design, was also involved in recruitment and therefore not blind to participant's group membership. While the possibility of bias should be considered, all of the tests used involved standardized administration and scoring procedures, and would not be expected to be affected.

This study attempted to address some of the problems inherent in the research in this area by conducting a small well-controlled study, including an anxious control group, excluding patients with illicit drug or high alcohol use, and excluding patients in any group who take antidepressant or other psychotropic medication. Furthermore, we present the results in a comprehensive manner that is amenable to inclusion in future meta-analyses. As a result of these strict inclusion criteria, the BZD group and the nonmedicated ANX control group proved to be very difficult to recruit and may have introduced the type of selection bias previously discussed. Recruitment difficulties also resulted in a smaller than anticipated sample size. Irrespective of these difficulties, the authors maintain that this approach to furthering our understanding of the nature of effects caused by long-term use of benzodiazepines is the most appropriate.

This study provides additional support for the hypothesis that long-term benzodiazepine impacts negatively on cognitive functioning in a number of areas. Given that the length of abstinence from benzodiazepines in this study was between 1 and 29 years, the persisting cognitive effects observed indicate that the impairments may also be long-term. These findings have important implications for those considering taking benzodiazepine medication, as well as those considering prescribing these drugs. Informed decisions regarding this therapy should be only made after considering the potential long-term impacts in conjunction with proposed treatment benefits.

ACKNOWLEDGMENTS

The authors would like to gratefully acknowledge Gwenda Cannard and the staff at TRANX for their assistance in the recruitment of previous long-term benzodiazepine users.

REFERENCES

Aranko, K., Mattila, M.J., & Seppala, T. (1983). Development of tolerance and cross-tolerance to the psychomotor actions of lorazepam and diazepam in man. *British Journal of Pharmacology*, *15*, 545–552.

Ashton, H. (1986). Adverse effects of prolonged benzodiazepine use. *Adverse Drug Reaction Bulletin*, *118*, 440–443.

Ashton, H. (1995). Toxicity and adverse consequences of benzodiazepine use. *Psychiatric Annals*, *25*, 158–165.

Australian Institute of Health and Welfare. (2000). *Australia's Health: The Seventh Biennial Health Report of the Australian Institute of Health and Welfare*. Canberra: Australian Institute of Health and Welfare. Cat. 19.

Balter, M.B., Manheimer, D.I., Mellinger, G.D., & Uhlenhuth, E.H. (1984). A cross-national comparison of anti-anxiety/sedative drug use. *Current Medical Research Opinion, 8 (Suppl. 4)*, 5–20.

Barker, M.J., Greenwood, K.M., Jackson, M., & Crowe, S.F. (2004a). The cognitive effects of long-term benzodiazepine use: A meta-analysis. *CNS Drugs*, *18*, 37–48.

Barker, M.J., Greenwood, K.M., Jackson, M., & Crowe, S.F. (2004b). Persistence of cognitive effects after withdrawal from long-term benzodiazepine use: A meta-analysis. *Archives of Clinical Neuropsychology*, *19*, 437–454.

Barker, M.J., Jackson, M., Greenwood, K.M., & Crowe, S.F. (2003). Cognitive effects of Benzodiazepine use: A review. *Australian Psychologist*, *38*, 202–231.

Benton, A.L., deS Hamsher, K., Varney, N.R., & Spreen, O. (1983). *Contributions to neuropsychological assessment*. New York: Oxford University Press.

Bergman, H., Borg, S., Engelbrektson, K., & Vikander, B. (1989). Dependence on sedative-hypnotics: Neuropsychological impairment, field dependence and clinical course in a 5-year follow-up study. *British Journal of Addiction*, *84*, 547–553.

Bergman, H., Borg, S., & Holm, S. (1980). Neuropsychological impairment and exclusive abuse of sedatives or hypnotics. *American Journal of Psychiatry*, *137*, 215–217.

Birzele, H.J. (1992). Benzodiazepine induced amnesia after long-term medication and during withdrawal. *European Review of Applied Psychology*, *42*, 277–282.

Brosan, L., Broadbent, D., Nutt, D., & Broadbent, M. (1986). Performance effects of diazepam during and after prolonged administration. *Psychological Medicine*, *16*, 561–571.

Chen, Y. (1990). Long-term benzodiazepine treatment: Is it ever justified? *Human Psychopharmacology Clinical and Experimental*, *5*, 301–312.

Cohen, J. (1988). *Statistical power analyses for the behavioral sciences* (2nd ed.). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Curran, H.V. (1986). Tranquillising memories: A review of the effects of benzodiazepines on human memory. *Biological Psychology*, *23*, 179–213.

Curran, H.V. (1992). Memory functions, alertness and mood of long-term benzodiazepine users: A preliminary investigation of the effects of a normal daily dose. *Journal of Psychopharmacology*, *6*, 69–75.

Curran, H.V., Bond, A., O'Sullivan, G., Bruce, M., Marks, I., Lelliot, P., Shine, P., & Lader, M. (1994). Memory functions, alprazolam and exposure therapy: A controlled longitudinal study of agoraphobia with panic disorder. *Psychological Medicine*, *24*, 969–976.

Golombok, S. (1989). Causes, effects and treatment of long-term benzodiazepine use: A review of psychological perspectives. *Human Psychopharmacology Clinical and Experimental*, *4*, 15–20.

Golombok, S., Moodley, P., & Lader, M. (1988). Cognitive impairment in long-term benzodiazepine users. *Psychological-Medicine*, *18*, 365–374.

Gorenstein, C., Bernik, M.A., & Pompeia, S. (1994). Differential acute psychomotor and cognitive effects of diazepam on long-term benzodiazepine users. *International Clinical Psychopharmacology*, *9*, 145–153.

Gorenstein, C., Bernik, M.A., Pompeia, S., & Marcourakis, T. (1995). Impairment of performance associated with long-term use of benzodiazepines. *Journal of Psychopharmacology*, *9*, 313–318.

Kaufman, A.S. & Kaufman, N.L. (1983). *K-ABC Kaufman Assessment Battery for Children*. Circle Pines, Minnesota: American Guidance Service.

Kirk, R.E. (1995). *Experimental design: Procedures for the behavioural sciences* (3rd ed.). Pacific Grove, California: Brooks/Cole Publishing Company.

Lezak, M.D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

Lucki, I., Rickels, K., & Geller, A.M. (1986). Chronic use of benzodiazepines and psychomotor and cognitive test performance. *Psychopharmacology*, *88*, 426–433.

Malec, J.F., Ivnik, R.J., Smith, G.E., Tangalos, E.G., Petersen, R.C., Kokmen, E., & Kurland, L.T. (1992). Visual Spatial Learning Test: normative data and further validation. *Psychological Assessment*, *4*, 433–441.

Massin-Krauss, M., Bacon, E., & Danion, J.M. (2002). Effects of the benzodiazepine lorazepam on monitoring and control processes in semantic memory. *Consciousness and Cognition*, *11*, 123–127.

McMillan, J. & Jones, F.L. (2000). The ANU3_2 scale: A revised occupational status for Australia. *Journal of Sociology*, *36*, 64–80.

Mellinger, G.D., Balter, B.B., & Uhlenhuth, E.H. (1984). Prevalence and correlates of the long-term regular use of anxiolytics. *Journal of the American Medical Association*, *251*, 375–379.

Mintzer, M.Z., Griffiths, R.R., Contoreggi, C., Kimes, A.S., London, E.D., & Ernst, M. (2001). Effects of Triazolam on brain activity during episodic memory encoding. *Neuropsychopharmacology*, *25*, 744–756.

Nelson, H.E. & Willison, J. (1991). *National Adult Reading Test (NART): Test Manual*. (2nd ed.). Windsor, UK: NFER Nelson.

Petursson, H., Gudjonsson, G.H., & Lader, M.H. (1983). Psychometric performance during withdrawal from long-term benzodiazepine treatment. *Psychopharmacology*, *81*, 345–349.

Reitan, R.M. & Wolfson, D. (1985). *The Halstead-Reitan Neuropsychological Test Battery*. Tucson, Arizona: Neuropsychology Press.

Rosenthal, R. (1991). *Meta-analytic procedures of social research* (revised ed.). London, UK: Sage Publications.

Sakol, M.S. & Power, K.G. (1988). The effects of long-term benzodiazepine treatment and graded withdrawal on psychometric performance. *Psychopharmacology*, *95*, 135–138.

Salzman, C., Fisher, J., Nobel, K., & Glassman, R. (1992). Cognitive improvement following benzodiazepine discontinuation in elderly nursing home residents. *International Journal of Geriatric Psychiatry*, *7*, 89–93.

Spielberger, C.D. (1983). *State-Trait Anxiety Inventory*. Palo Alto, California: Mind Garden.

Spreen, O. & Strauss, E. (1998). *A compendium of neuropsychological tests: Administration, norms, and commentary* (2nd ed.). New York: Oxford University Press.

Tata, P.R., Rollings, J., Collins, M., Pickering, A., & Jacobson, R.R. (1994). Lack of cognitive recovery following withdrawal from long-term benzodiazepine use. *Psychological Medicine*, *24*, 203–213.

Tiffin, J. (1987). *Purdue Pegboard Examiner's Manual* (3 ed.). Illinois: National Computer Systems.

Toenne, U., Hiltunen, A.J., Vikander, B., Engelbrektsson, K., Bergman, H., Bergman I., Leifman, H., & Borg, S. (1995). Neuropsychological changes during steady-state drug use, withdrawal and abstinence in primary benzodiazepine-dependent patients. *Acta Psychiatrica Scandinavica*, *91*, 299–304.

Walsh, K. (1985). *Understanding brain damage: A primer of neuropsychological evaluation*. London, UK: Churchill Livingstone.

Wechsler, D. (1997). *Wechsler Memory Scale–Third Edition (WMS–III)*. San Antonio, Texas: The Psychological Corporation.

Wechsler, D. (1999). *Wechsler Abbreviated Test of Intelligence (WASI)*. San Antonio, Texas: The Psychological Corporation.

Wolkowitz, O.M., Weingarter, H., Thompson, K., Pickar, D., Paul, S.M., & Hommer, D.W. (1987). Diazepam-induced amnesia: A neuropharmacological model of an 'organic-amnesic syndrome'. *American Journal of Psychiatry*, *144*, 25–29.

Zandstra, S.M., Furer, J.W., van de Lisdonk, E.H., van't Hof, M., Bor, J.H.J., van Weel, C., & Zitman, F.G. (2002). Different study criteria affect the prevalence of benzodiazepine use. *Social Psychiatry and Psychiatric Epidemiology*, *37*, 139–144.