Behavioral Pharmacology and Verbal Behavior: Diazepam Effects on Verbal Self-Reports

Thomas S. Critchfield
Auburn University

Diazepam (10 mg) was administered to two men performing a delayed matching-to-sample task in which the number of elements in a compound sample stimulus (one of which appeared among 4 comparison stimuli) was manipulated from 1 to 3. After each trial, subjects pressed either a "Yes" or "No" button in response to a computer-presented query about whether the last choice met a point contingency requiring selection of the matching comparison stimulus within a time limit. Diazepam simultaneously produced marginal decreases in matching-to-sample performance and more pronounced decreases in the accuracy of self-reports about the same performance. Diazepam selectively increased false reports of success; false reports of failure were not systematically affected. A signal-detection analysis summarized these patterns as a decrease in self-report discriminability (A') with no systematic change in bias ($B'H$). These preliminary results converge with those of clinical lore and the results of studies with other benzodiazepine drugs in suggesting that diazepam can produce an "overconfidence" in performance self-evaluation, the mechanisms and parameters of which remain to be identified. The results were inconsistent with those of one previous study of diazepam's effects on performance self-evaluation, but given procedural differences between the two studies, the discrepancy may reflect the functional independence of verbal operant classes in Skinner's (1957) taxonomy.

Skinner's (1957) conceptual analysis of verbal behavior focused on the functions of verbal responses in a context defined by operant consequences, establishing operations, and discriminative stimuli. But to be useful an analysis of verbal behavior need not employ Skinner's original categories of verbal operants (e.g., Skinner, 1989; Zettle & Hayes, 1982), or even deal exclusively with operant processes. Pharmacological effects, for example, fall outside the usual Skinnerian operant formulation but, like other non-operant phenomena, should be of interest in a thoroughgoing analysis because they may set boundaries within which operant effects may occur, or interact with operant control of verbal behavior (e.g., Poling &

LeSage, 1992). Although scant literature describes pharmacological effects on verbal behavior, a few such effects have been identified. For example, the stimulant d-amphetamine appears to increase free-operant rates of speaking (Griffiths, Stitzer, Corker, Bigelow, & Liebson, 1977; Stitzer, Griffiths, & Liebson, 1978), and to increase the relative reinforcing potency of verbal-social interaction (Higgins, Hughes, & Bickel, 1989; Higgins & Stitzer, 1988).

Pharmacological effects on verbal behavior should be of special interest in the case of verbal self-reports, in part because self-reports comprise such an important part of everyday interactions (Skinner, 1957). Additionally, as a type of self-discrimination, self-reports may be functionally similar to the kinds of ongoing, covert self-observation responses many scientists believe provide the basis for integrated, dynamic performances such as driving an automobile or having a conversation (e.g., Carver & Scheier, 1981; Rabbitt, 1979; Skinner, 1957). In this sense, it would be

This study was conducted while the author was a postdoctoral fellow under Grant 03389 from the National Institute on Drug Abuse to Roland R. Griffiths, whose support of the research is much appreciated. John Texter capably collected the data and helped prepare them for analysis. I thank Mark Sundberg and two anonymous reviewers for helpful comments. Address correspondence to the author at the Department of Psychology, Auburn University, Auburn, AL 36849.

reasonable to view verbal self-reports as a possible tool for assessing drug-induced changes in "judgment."

Conventional wisdom holds that, when intoxicated, individuals may evaluate their behavior imperfectly and thus, for example, fail to notice and correct for performance errors. Yet the empirical literature provides little support for the conventional wisdom regarding drug effects on judgment. A recent review of the literature found that few relevant studies have been conducted, and that existing studies often have employed questionable methods and produced conflicting results. The review suggested that, conventional wisdom notwithstanding, there is virtually no unambiguous evidence that drugs impair performance self-evaluation (Critchfield & Schlund, 1992).

The present study was a preliminary investigation of the effects of diazepam (Valium®) on verbal self-reports about performance, using a procedure shown previously to be sensitive to the effects of several non-pharmacological manipulations (e.g., Critchfield, 1993; Critchfield & Perone, 1993). The study compared self-reports under the influence of placebo and 10 mg diazepam. Diazepam was chosen because, along with several other benzodiazepines which also are used as anxiolytics, it is among the more commonly prescribed psychoactive drugs; 10 mg is a common therapeutic dose for outpatient treatment of anxiety (Baldessarini, 1990). Thus, large numbers of patients perform everyday activities, including driving automobiles and operating machinery, under the influence of this dose. Diazepam was also of interest as a member of the benzodiazepine class. Studies suggest that, compared to other sedative drugs, the benzodiazepines triazolam and lorazepam are especially likely to inflate self-estimates of performance (Roache & Griffiths, 1985, 1987b).

Diazepam's effects on performance self-evaluation apparently have been evaluated in only one study to date. Subjects attempted to recall 12 previously-studied words, and rated their confidence that each word they produced during the recall test

actually was among those studied (Roy-Byrne, Uhde, Holcomb, Thompson, King, & Weingartner, 1987). Diazepam (10 mg) neither impaired free recall nor altered average confidence ratings, prompting the authors to conclude that accurate self-monitoring remained intact. Two factors may promote skepticism, however, about this conclusion. First, the conclusion apparently was bolstered by a positive correlation between individual subjects' ratings of sedative symptoms and performance on ancillary memory tests – neither of which was directly related to the free recall which subjects self-evaluated. Second, confidence ratings apparently did not track a non-significant trend toward diazepam-induced impairment on the free recall task, suggesting that the ratings were insensitive to changes in performance – an effect consistent with the notion that diazepam disrupts performance self-evaluation. It is possible, for example, that with more than 12 words in the test battery, or more assessments per individual, that diazepam effects on self-evaluation might have been detected. Alternatively, it is possible that changes in individual self-evaluation were lost in the averaging necessary for statistical analysis.

In the present experiment, subjects made self-reports about their performance in a delayed matching-to-sample (DMTS) task in which correct choices, made within approximately one-half second of the appearance of comparison stimuli, earned points. Immediately following each choice, subjects responded "Yes" or "No" to a computer-presented query asking whether the last DMTS choice met the point contingency. Unlike the one previous study using diazepam (Roy-Byrne, et al., 1987), self-reports were assessed across hundreds of trials, permitting ample opportunity to evaluate patterns of self-reporting in individual subjects.

This study also employed an analytical strategy in which self-reports about DMTS performance were viewed in the context of a contingency matrix, shown in Figure 1, involving two levels of DMTS outcomes (success and failure) and two levels of self-

	Successful	Unsuccessful
"I succeeded"	CORRECT DETECTION	FALSE DETECTION
"I failed"	FALSE REJECTION	CORRECT REJECTION

Fig. 1. Contingency matrix showing self-reports of delayed matching-to-sample success as a function of actual success; cells are labelled using response categories derived from signal-detection theory.

reports ("I succeeded" and "I failed"). Conceivably, a drug could disrupt performance self-evaluation by selectively increasing rates of *false detections* (reports of success following DMTS failure), selectively increasing rates of *false rejections* (reports of failure following DMTS success), or nonselectively increasing both types of self-report errors.

METHOD

Subjects

Two men, 33 and 31 years old, volunteered as paid participants in a study on "Drug Effects on Performance and Judgment." A preliminary interview and physical screening indicated normal vital signs and no significant medical or psychiatric symptoms. Both men smoked tobacco cigarettes and reported using alcohol occasionally, but no illicit drugs. Neither was taking prescription medication. The men provided written informed consent before participating.

The men were asked not to consume solid food or caffeinated beverages for 12 hr before the start of an experimental session. Urine samples collected before each session were tested for the presence of illicit drugs (barbiturates, benzodiazepines, cocaine, methadone, and opiates) using an EMIT system (Syva Co., Palo Alto, CA) to ensure that diazepam was not administered to subjects already under the influence of these psychoactive substances. Breathalyzer tests assured that subjects

were negative for the presence of blood alcohol.

The men were paid an hourly wage of \$4.00 supplemented by earnings contingent on performance during experimental sessions. The men received \$10.00 at the end of each visit to the laboratory, and the remainder of their earnings at the end of their participation. On average, earnings totalled about \$7.00 to \$8.00 per hour.

Setting and Apparatus

The large room used for the study contained several small test booths, a desk and chair for a technician, a social area (sofa, table, chairs, television, and reading materials), and a kitchen area containing a refrigerator, toaster, and coffee maker. One test booth housed a table, chair, and a console (described by Critchfield & Perone, 1990), atop of which rested a monochrome video monitor. The console's sloping front panel contained four round, illuminable buttons (used in the DMTS task) arranged horizontally along the front, near the subject, and two small white lamps toward the back, near the monitor. Mounted on a small metal box attached at table level to each side of the console was a pushbutton (used to make self-reports). A microcomputer in an adjacent booth was used to control the experimental events and collect the data.

General Procedure

About one hour before drug was administered, urine and breathalyzer tests were conducted and the men ate a snack consisting of unbuttered toast, jelly, and juice. Pre-drug performance measures were obtained during the ensuing hour. Following drug administration, data were collected across a 3.5-hr time course. A lunch (sandwich, potato chips, and noncaffeinated soda or juice) was served following the 2.5-hr assessments. The men were free to talk, watch television, or read during the brief intervals during which data were not being collected.

To reduce the possibility of cumulative or carry-over drug effects, at least one placebo session separated diazepam

administrations (the interval between diazepam administrations was always at least three days). The informed consent agreement stated that the men might receive various drugs during their participation, including placebo, stimulants, sedatives, neuroleptics, minor tranquilizers, and muscle relaxants. Otherwise, subjects were blind to the type of drug and the schedule of drug administration. The consent agreement also stated that the purpose of the study was to see "how different drugs affect mood and behavior." At times specified below, the subjects read written instructions describing the operation of the console.

Drugs

During the study the men received 3 exposures each of placebo and 10 mg diazepam. The drugs were administered orally in capsules using a double-blind procedure in which neither the subjects nor the technician who conducted sessions were aware of drug or drug class under investigation or the contents of the capsules on any given day.

Self-Reports of DMTS Performance.

Typical DMTS trial. The procedure was similar to one described previously (Critchfield & Perone, 1993). Each man sat alone in the self-report booth facing the console and monitor. The monitor screen was divided into two portions, each surrounded by a distinctive border. At the start of each trial the four round buttons at the front of the console became illuminated, and the monitor displayed the message, "HOLD LIGHTED BUTTONS DOWN" in the top (DMTS) portion of the screen. Depressing all four buttons simultaneously (subjects typically used the first and fourth fingers of each hand) produced a 700-ms display of 1, 2, or 3 sample stimuli, followed by a 1-s retention interval during which the screen was blank. Four horizontally-aligned comparison stimuli, one of which matched a sample shape, then appeared in locations corresponding to the four illuminated buttons. Release of the button corresponding to the matching

comparison stimulus was counted as correct. When a correct response occurred within a time limit, the subject earned points worth money (1 point = 1 cent). For S1 and S2, time limits were 575 ms and 500 ms, respectively, after comparison stimuli appeared [these values were selected on the basis of preliminary training performance to produce intermediate levels of DMTS success]. No immediate stimulus change indicated when the time limit had elapsed. Trials were separated by an inter-trial interval (ITI) lasting about 1 s, during which all areas of the screen were cleared. A variety of error messages discouraged responses that deviated from this procedure (see Critchfield & Perone, 1990).

Each sample and comparison stimulus consisted of a 6 by 3 matrix of rectangular cells, as few as 2 or as many as 18 of which could be illuminated, to produce a shape approximately 10 by 13 mm in size (see Critchfield, 1993). The stimuli on each trial were drawn randomly without replacement from a pool of several thousand unique shapes, with the obvious exception that one sample stimulus always matched one comparison stimulus. The number of sample stimuli (1, 2, or 3), the sample stimulus to be reproduced among the comparisons (on 2-sample trials and 3-sample trials), and the location in the comparison-stimulus display where the matching stimulus appeared, all varied quasi-randomly across trials.

Self-reports. When scheduled, self-reports took place immediately following each DMTS response. The top, DMTS portion of the screen was cleared and the bottom, self-report portion displayed the query, "Did you score?" (as described below, this was the wording used during preliminary training to indicate when a response qualified for reinforcement). Below it were the messages "<- Yes," printed 3 cm from the left edge of the self-report area of the screen, and "No ->," printed 3 cm from the right edge. Pressing the button attached to the console's left side registered a "Yes" report and pressing the button attached to the console's right side registered a "No" report. Other responses were ineffective.

Self-reports were followed by feedback messages (described below), when scheduled.

Sessions. Each daily session consisted of nine work periods. The first two occurred before drug administration (their data are pooled as the pre-drug measure in the Results) and the remainder at 0.5-hr intervals thereafter. Each work period consisted of 60 trials, 20 with each type of DMTS trial (1, 2, or 3 samples). A 20-s intermission divided the work period into 30-trial blocks containing 10 trials of each type. Work periods typically took about 5 to 10 minutes to complete.

Preliminary training. Placebo was administered on four training days which preceded the actual experimental sessions. On Day 1, the men began practicing the DMTS task after reading the printed instructions reproduced in the first part of the Appendix. In addition, a card in the booth stated that "Points today are earned from Matching. 1 point = 1 cent." No self-reports took place. All parameters were as described above except that the time limit on DMTS responding began at 3000 ms for the first 30 trials, and decreased with each subsequent block of 30 trials as follows: 2500, 2000, 1750, 1500, 1250, 1000, 900, 800, 800, 750, 750, 650, 650, and finally the terminal value of the time limit. Thus, by the end of the first day the time limit had reached its final value).

On each trial, choice of a comparison stimulus produced a 1.5-s display of three feedback messages accompanied by illumination of the small white lights at the top of the console's work panel. The first message indicated whether the choice of a comparison stimulus had been "Correct" or "Wrong." The second stated whether the choice had occurred "Fast Enough" or "Too Slow." If the choice was both correct and timely, a third message stated, "You Scored! 2 points have been added to your total." Otherwise, the third message stated, "No Score. No change in your total." The procedures on Day 2 replicated those at the conclusion of Day 1 except that the DMTS time limit remained constant and the feedback message was simplified. If selection

of a comparison stimulus was both correct and timely, the message "You Scored!" appeared for 1 s. Otherwise, the message, "No Score" appeared for 1 s. No specific information about speed and accuracy was provided, and no lights accompanied the feedback messages.

On Day 3 the self-report procedure was introduced. Subjects read the printed instructions reproduced in the second part of the Appendix. Feedback about DMTS success no longer followed selection of a comparison stimulus. Instead, the self-report query ("Did you score?") ensued. Emission of a self-report produced a 1-s feedback message ("Results of your report" followed by "Correct - 1 point bonus" or "Wrong - 1 point penalty") and then the ITI. The conditions on Day 4 replicated those of Day 3 with the exception that consequences for self-report accuracy were delivered on a random-ratio 3 schedule. When no consequences were scheduled, completion of the self-report led directly to the ITI. When consequences occurred, a feedback message following the self-report stated Results of your report" followed by "Correct - 3 point bonus" or "Wrong - 3 point penalty." Conditions in the main study were identical to those on Day 4 of pre-training.

RESULTS

Data from the multiple determinations of placebo and diazepam were summed in order to provide a sufficiently large pool of trials for analysis. This was necessary because the 60 trials which could be completed at each time point during a session, divided by three levels of DMTS sample-stimulus complexity and the four response categories of signal detection analysis shown in Figure 1, provided too few data for separate analysis.

Figure 2 shows the percentage of successful DMTS responses (those that were both accurate and faster than the time limit) across a 3.5-hr time course. For both subjects, success rates on placebo days (open circles) tended to decrease as the number of DMTS sample stimuli increased. Relative to placebo days,

Fig. 2. Percent successful delayed matching-to-sample responses, across a 3.5-hr time course after drug administration.

diazepam tended to decrease success rates (filled circles): Modest decrements occurred by 1 to 2 hr post-drug for both subjects on 1-sample trials and 2-sample trials, but only for S1 on 3-sample trials.

Figure 3 provides a broad summary of the effects of diazepam on self-reports about DMTS performance, showing self-report accuracy for 1-sample, 2-sample, and 3-sample trials, across a 3.5-hr time course. Compared to placebo, diazepam decreased self-report accuracy on all three types of trials by 1 to 2 hr post-drug. Figure 4 examines the accuracy decrement more closely, showing rates of the two categories of inaccurate self-reports shown in Figure 1, false detections (reports of "success" following unsuccessful DMTS responses) and false rejections (reports of "failure" following successful DMTS responses).

The top two rows of panels in Figure 4 show rates of false detections (false detections divided by the sum of false plus correct detections) across a 3.5-hr time course. For both subjects, diazepam (filled circles) tended to increase false detection rates above those on placebo days (open circles), and the effect was more pronounced on 1-sample and 2-sample trials than on 3-sample trials. The bottom two rows of panels in Figure 4 show rates of false rejections (false rejections divided by the sum of false plus correct rejections), which were low on placebo days, compared to rates of false detections, and not systematically affected by diazepam.

The determination of rates of false detections and false rejections permit indices of discriminability and bias, as formulated for signal-detection analysis (Green & Swets, 1966), to be applied to the analysis of self-reports (Critchfield, 1993; Critchfield &

Fig. 3. Percent accurate self-reports, across a 3.5-hr time course following drug administration.

Perone, 1993). Figure 5 shows peak self-report discriminability (A') and bias ($B'H$) scores, calculated using nonparametric indices described by Grier (1971). Because diazepam primarily affected rates of false detections, peak effects here reflect the time point of greatest deviation from pre-drug baseline of false detection rates.

In the present context, the A' measure of detectability estimates an individual's ability to recognize a "signal" consisting of successful DMTS responses. Scores can range from 0.0 to 1.0, with 0.5 representing chance reporting and higher scores representing better detection. Compared to placebo, diazepam reduced self-report discriminability; for both subjects, the magnitude of difference was least for 3-sample trials.

In the present context, the $B'H$ measure of bias estimates an individual's tendency to report DMTS success or failure, regardless of the actual status of the DMTS

response. Scores can range from -1.0 to +1.0, with negative scores indicating a bias for reporting success and positive scores indicating a bias for reporting failure. Bias scores always were negative, indicating a consistent bias for reporting successful DMTS responses. Diazepam had no systematic effect on this pattern.

DISCUSSION

A low (therapeutic-level) dose of diazepam tended to produce modest decreases in DMTS success and simultaneously decreased the accuracy of self-reports about DMTS performance. A signal-detection analysis proved useful in characterizing the effects of diazepam more precisely than could be accomplished using the more global measure of accuracy. Decrements in self-report accuracy reflected increases in false detection rates with no corresponding change in false

Fig. 4. Rates, per opportunity, of two types of self-report errors, false detections and false rejections, across a 3.5-hr time course after drug administration.

rejection rates, resulting in a decrease in self-report discriminability but no systematic change in self-report bias. These findings are orderly enough to serve as food for thought about the "verbal behavioral pharmacology" of diazepam (see

below), but must await verification through more systematic research. In particular, future studies will need to test a range of diazepam doses, to determine whether effects are dose-dependent; and to compare diazepam with a positive control

Fig. 5. Peak self-report discriminability and bias as a function of the number of sample stimuli on each delayed matching-to-sample trial.

drug with similar sedative effects (e.g., barbiturates), to determine whether different drugs produce unique profiles of disruption in performance self-evaluation (e.g., see Evans, Critchfield, & Griffiths, 1992).

The results of the present study are consistent with anecdote and clinical lore in suggesting that diazepam can disrupt performance self-evaluation. In addition, the present results may help to resolve an ambiguity regarding two previous studies which appear to show that other benzodiazepines have similar effects. Roache and Griffiths (1985, 1987b) found that triazolam and lorazepam, compared to non-benzodiazepine sedatives, selectively inflated self-estimates of performance on psychomotor tasks lasting 60 to 90 s. Because benzodiazepine drugs impair remembering (Curren, 1986; Hindmarch & Ott, 1988) – possibly within seconds after presentation of information to be remembered (Roache & Griffiths, 1987a) – it remains unclear whether the pattern described by Roache and Griffiths reflects a specific failure to detect performance errors or simply a global memory deficit which manifested over the minute or more of each task's duration. Intoxicated subjects could either have failed to discriminate their performances, or instantly discriminated every

response but forgotten about some by the time they made their self-evaluations. Inaccurate self-evaluations would result in either case, but impairment of moment-by-moment self-monitoring abilities would have greater implications for everyday activities such as driving an automobile. Presumably, the present study, in which self-reports occurred immediately following referent responses, avoids the ambiguity of Roache and Griffiths's procedure, and suggests more directly that diazepam can impair moment-by-moment self-discriminations.

Of special interest in the present results is the fact that diazepam produced highly specific increases in false detections, rather than a global pattern of disruption of self-reports – producing what might be characterized as a state of drug-induced "over-confidence." Such a finding, if verified by further research, would be no less for important for conforming to conventional wisdom, given the scope of social problems created by drug use and abuse. Surprisingly little is known about how (or, indeed, whether) drugs influence "judgment." To date, inferences about the effects of drugs on performance self-evaluation have been based on correlations between performance measures and self-estimates of blood alcohol content or of general drug intoxication (e.g., Vogel-Sprott, 1975). Yet judgements about general drunkenness do not necessarily reflect *performance* self-evaluation.

Few drugs have been evaluated systematically enough, using direct measures of performance self-evaluation, to permit strong conclusions about their effects (Critchfield & Schlund, 1992). Drugs that have been studied frequently, such as alcohol, appear to have very complex profiles of action. For example, at blood concentrations of around .10%, alcohol has been found to impair self-evaluation of performance on a pursuit rotor task, but not of performance on a memory task or a digit-symbol substitution task (Lubin, 1979); to impair self-evaluation of flight-simulator performance by younger, but not older, pilots (Morrow, Leirer, Yesavage, &

Tinklenberg, 1991); and to increase the time needed to self-evaluate choice latency without impairing the quality of self-evaluation (Maylor, Rabbitt, & Connolly, 1989).

It seems safe to conclude that much remains to be learned about the effects of drugs on performance self-evaluation. In this regard, it is noteworthy that the present results appear to run counter to those of the one previous published study examining diazepam's effects on performance self-evaluation (Roy-Byrne, et al., 1987), which found no diazepam-associated impairment of self-reports about memory for a list of words. The Introduction to this report highlighted methodological differences between the two studies, but it is also interesting to note that, functionally, the self-reports of the two studies probably constitute different categories in Skinner's (1957) taxonomy of verbal operants. Self-reports in the Roy-Byrne, et al. (1987) study – the delayed repetition of a list of words – probably consisted of intraverbal relations. Self-reports in the present study, by virtue of their close contiguity to referent behavior, probably were tacts with some autoclitic and intraverbal properties. In this sense, the present results suggest that Skinner's taxonomy may be valuable in categorizing pharmacological effects on verbal behavior. Such an assumption is bolstered by the fact that benzodiazepine drugs have already been shown to differentially affect memory performances (Taylor & Tinklenberg, 1987) that would be differentially classified in Skinner's taxonomy.

If, as some have concluded, the experimental analysis of verbal behavior remains in its infancy (e.g., Oah & Dickinson, 1989; see also Dougherty, Nedelman, & Alfred, 1993), then a systematic experimental analysis of the relations between drugs and verbal behavior can be described as still waiting to be born. Ideally, research on drug-behavior interactions can teach us about both drugs and behavior (Branch, 1984). Similarly, pharmacological effects are not considered in Skinner's (1957) conceptual analysis of verbal behavior, but

if the analysis truly provides a general purpose account of verbal phenomena then its application to problems in behavioral pharmacology should advance both fields. In this regard, a Skinnerian interpretation of the contrast between the present results and those of Roy-Byrne et al. (1987) provides some reason for optimism, but establishes nothing with empirical certainty (e.g., procedures of the two studies can just as easily be distinguished using constructs from cognitive information-processing theory). Until behavioral pharmacology studies are designed explicitly to explore the value of Skinner's analysis – something not undertaken in the present investigation – the value of *Verbal Behavior* to behavioral pharmacology will remain a matter of speculation (e.g., Poling & LaSage, 1992).

REFERENCES

- Baldessarini, R. J. (1990). Drugs and the treatment of psychiatric disorders. In A. G. Gilman, T. W. Rall, A. S. Nies, & P. Taylor (Eds.), *The pharmacological basis of therapeutics* (8th ed.) (pp. 383-435). New York: Pergamon.
- Branch, M. N. (1984). Rate dependency, behavioral mechanisms, and behavioral pharmacology. *Journal of the Experimental Analysis of Behavior*, 42, 511-522.
- Carver, C. S., & Scheier, M. F. (1981). *Attention and self-regulation: A control-theory approach to human behavior*. New York: Springer-Verlag.
- Critchfield, T. S. (1993). Signal detection properties of verbal self-reports. *Journal of the Experimental Analysis of Behavior*, 60, 495-514.
- Critchfield, T. S., & Perone, M. (1990). Verbal self-reports of delayed matching to sample by humans. *Journal of the Experimental Analysis of Behavior*, 53, 321-344.
- Critchfield, T. S., & Perone, M. (1993). Verbal self-reports about matching-to-sample success: Effects of the number of elements in a compound sample stimulus. *Journal of the Experimental Analysis of Behavior*, 59, 193-214.
- Critchfield, T. S., & Schlund, M. (1992, May). Do drugs really impair judgment? Drug effects on performance self-evaluation. Paper presented at the 18th Association for Behavior Analysis Convention, San Francisco, CA.
- Curren, H. V. (1986). Tranquillising memories: A review of the effects of benzodiazepines on human memory. *Biopsychology*, 23, 179-213.
- Dougherty, D. M., Nedelmann, M., & Alfred, M. (1993). An analysis and topical bibliography of the last ten years of human operant behavior: From minority to near majority (1982-1992). *Psychological Record*, 43, 501-530.
- Evans, S. M., Critchfield, T. S., & Griffiths, R. R. (1992). Abuse liability assessment of anxiolytic/hypnotics: Rationale and laboratory lore. *British Journal of Addiction*, 86, 1625-1632.

- Evans, S. M., Funderburke, F. R., & Griffiths, R. R. (1990). Zolpidem and triazolam in humans: Behavioral and subjective effects and abuse liability. *Journal of Pharmacology and Experimental Therapeutics*, 255, 1246-1255.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin*, 75, 424-429.
- Griffiths, R. R., Stitzer, M., Corker, K., Bigelow, G., & Liebson, I. (1977). Drug-produced changes in human social behavior: Facilitation by d-amphetamine. *Pharmacology Biochemistry & Behavior*, 7, 365-372.
- Higgins, S. T., Hughes, J. R., & Bickel, W. K. (1989). Effects of d-amphetamine in choice of social versus monetary reinforcement: A discrete-trial test. *Pharmacology Biochemistry & Behavior*, 34, 297-301.
- Higgins, S. T., & Stitzer, M. L. (1988). Time allocation in a concurrent schedule of social interaction and monetary reinforcement: Effects of d-amphetamine. *Pharmacology Biochemistry & Behavior*, 31, 227-231.
- Hindmarch, I., & Ott, H. (Eds.). (1988). *Benzodiazepine receptor ligands, memory, and information processing*. New York: Springer-Verlag.
- Lubin, R. A. (1979). Influences on alcohol, interpersonal feedback, and drinking experience upon performance and judgment. *Perceptual and Motor Skills*, 48, 95-104.
- Maylor, E. A., Rabbitt, P. M. A., & Connolly, S. A. V. (1989). Rate of processing and judgment of response speed: Comparing the effects of alcohol and practice. *Perception & Psychophysics*, 45, 431-438.
- Morrow, D., Leirer, V., Yesavage, J., & Tinklenberg, J. (1991). Alcohol, age, and piloting: Judgment, mood, and actual performance. *International Journal of the Addictions*, 26, 669-683.
- Oah, S., & Dickinson, A. M. (1989). A review of empirical studies of verbal behavior. *Analysis of Verbal Behavior*, 7, 53-68.
- Poling, A., & LaSage, M. (1992). Rule-governed behavior and behavioral pharmacology: A brief comment on an important topic. *Analysis of Verbal Behavior*, 10, 37-44.
- Rabbitt, P. M. A. (1979). Current paradigms and models in human information processing. In V. H. Hamilton, & D. M. Warburton (Eds.), *Human stress and cognition: An information processing approach* (pp. 115-140). New York: Wiley.
- Roache, J., & Griffiths, R. R. (1985). Comparison of triazolam and pentobarbital: Performance impairment, subjective effects and abuse liability. *Journal of Pharmacology and Experimental Therapeutics*, 234, 120-133.
- Roache, J., & Griffiths, R. R. (1987a). Interactions of diazepam and caffeine: Behavioral and subjective dose effects in humans. *Pharmacology, Biochemistry, & Behavior*, 26, 801-812.
- Roache, J., & Griffiths, R. R. (1987b). Lorazepam and meprobamate dose effects in humans: Behavioral effects and abuse liability. *Journal of Pharmacology and Experimental Therapeutics*, 243, 978-988.
- Roy-Byrne, P. P., Uhde, T. W., Holcomb, H., Thompson, K., King, A. K., & Weingartner, H. (1987). Effects of diazepam on cognitive processes in normal subjects. *Psychopharmacology*, 91, 30-33.
- Schutz, H. (1982). *Benzodiazepines: A handbook*. New York: Springer-Verlag.
- Skinner, B. F. (1957). *Verbal behavior*. Englewood Cliffs, NJ: Prentice-Hall.
- Skinner, B. F. (1989). The listener. In B. F. Skinner, *Recent issues in the analysis of behavior* (pp. 35-47). Columbus, OH: Merrill.
- Stitzer, M. L., Griffiths, R. R., & Liebson, I. (1978). Effects of d-amphetamine on speaking in isolated humans. *Pharmacology Biochemistry & Behavior*, 9, 57-63.
- Taylor, J. L., & Tinklenberg, J. R. (1987). Cognitive impairment and benzodiazepines. In H. Y. Meltzer (Ed.), *Psychopharmacology: The third generation of progress* (pp. 1449-1454). New York: Raven.
- Vogel-Sprott, M. (1975). Self-evaluation of performance and the ability to discriminate blood alcohol concentrations. *Journal of Studies on Alcohol*, 36, 1-10.
- Zettle, R. D., & Hayes, S. C. (1982). Rule-governed behavior: A potential theoretical framework for cognitive-behavior therapy. In P. C. Kendall (Ed.), *Advances in cognitive-behavioral research and therapy* (pp. 73-118). New York: Academic Press.

APPENDIX

Instructions Read Before Day 1 of Preliminary Training

In front of you is a console containing several lights and buttons, plus a video screen. Your job is to make decisions based on information presented on your screen, and to indicate your decisions using buttons on the console. Each time you work you will be given several decision-making opportunities (called a "trial"). At the start of each trial your screen will ask you to "HOLD DOWN THE LIGHTED BUTTONS." When you depress these four buttons, one or more "sample" shapes will appear on your screen, then disappear. Shortly thereafter, four "test" shapes will appear. Your job is to indicate which of these test shapes matches one of the samples. You can do so by releasing the lighted button corresponding to the matching shape. Note that you must hold down the lighted buttons until you are ready to make your decision. If you release too soon, the trial will cancel and start again, wasting time in which you could be earning points. You can earn a point each time you choose the correct (matching) test shape. In order to earn a point, your decision must be both correct, and within a time limit. That is, you may have to release the correct side button rather quickly to earn

a point. To start with, you will have a relatively long amount of time to make each decision. Later on, the time limit will become more stringent. Do the best you can under the time constraints. Today, after you complete each trial, messages on your screen will tell you whether you earned a point. Later on, you may be given less information about your decisions. Please complete all the trials each time you sit down at the console. Do not ask questions or leave the room during a work period. To minimize distractions while you are working, we ask that you wear the headphones on the table in front of you. Beyond the information contained in these instructions, it is up to you to decide how to operate the console to your best advantage. These instructions contain all the information we can provide at this time. If you have any questions, please ask them now.

Instructions Read Before Day 3 of Preliminary Training

Starting with your next session, the way you earn points will change in two ways. First, each time you make a correct choice within the time limit, you will still “score.” However, each score is now worth 1 point instead of 2 points. In addition, you will no longer receive direct feedback about the success of your choices. Second, after each choice, the computer will ask you whether you believe you scored. If your answer is accurate, you will earn a 1 point bonus. If not, a 1 point penalty will be subtracted from your total. [Note that you can still determine the success of your choices. For example, if you report that you scored, and your report turns out to be correct, then you know that, in fact, you scored.] To summarize, each trial will still be worth a possible 2 points. The change is that half comes from the matching task, and half now comes from your reports about your performance.