A test of sex differences in instrumental appetitive learning by rats¹

MELVIN H. MARX, UNIVERSITY OF MISSOURI, Columbia, Missouri 65201

A total of 48 laboratory rats, half male and half female, were trained and extinguished on instrumental bar-pressing for either 16% or 64% sucrose incentives. No consistent sex differences were found in either acquisition or variability of bar-press responding, or in extinction. The only statistically reliable difference indicated that males were more variable in responding for the 16% sucrose incentive, as measured by the range of scores over each week. It is concluded that no real sex differences in learning proficiency obtain under the experimental conditions employed, and that therefore the practice of using mainly or exclusively female rats in our laboratory is methodologically sound.

This experiment was designed to determine if female rats, with their 4- to 5-day estrous cycle, are more variable than male rats in the acquisition and extinction of a bar-press response. Apart from its contribution to the problem of sex differences per se, the experiment was designed to support or question the policy in my laboratory over the past several years of using mainly, and usually exclusively, female rats. This policy was originally adopted because females seemed to be more active as well as more tractable than males, especially as the animals grew past adolescence (from the 80- to 90-day-old range at which they are usually started in experiments) and because we did not wish to mix the sexes and run the risk of contamination of behavior by sex odors and the like. The relevance of the latter precaution has recently been suggested by demonstrations that rats in linear apparatuses do follow odor trails of various sorts (e.g., Courtney, Reid, & Wasden, 1968; Davis & Ludvigson, 1969; Lester, 1968; Ludvigson & Sytsma, 1967; McHose & Ludvigson, 1966).

The use of females alone had been based upon the belief that they are at least as good Ss as males, and upon the assumption that their estrous cycle does not interfere with their performance, especially their bar-press performance in automated equipment of the sort that we have been using (cf. Marx et al, 1965, for a description). Nevertheless, this assumption has not been tested empirically and it was the purpose of the present experiment to do so. The advisability of doing so has recently been confirmed (following design and completion of the present experiment) by reports that 1-h samples of running behavior can detect the cycle (Hitt, Gerall, & Giantonio, 1968) and that female rats do show 4- to 5-day cyclic variations in duration of bar-press behavior (Moss, 1968). However, no test has been reported of a direct comparison of male and female rats in the kind of apparatus that we have been using and extrapolations from these other kinds of situations cannot safely be made. Sucrose concentration of the incentives was also manipulated in the experiment because it was felt that this variable might differentiate the sexes.

METHOD

Subjects

The Ss were 24 male and 24 female hooded rats (Long-Evans) from the laboratory colony from which most of our Ss have been drawn over the past several years. They were approximately 60 days old at the beginning of the experiment. Each sex was randomly divided into two groups of 12 Ss each prior to the initiation of the experiment.

Apparatus

The apparatus was a set of eight automated controlled-operant conditioning boxes with retractable bars and plate-type feeding magazines. A standard Hoeltge HB 11-A living cage was housed in a ventilated, sound-proofed wooden hull. This apparatus has been previously described in detail (Marx et al, 1965).

Design

The design was a simple 2 by 2 factorial, in which each sex was trained on either 16% or 64% sucrose concentration. The four groups are designated as M-16, M-64, F-16, and F-64. The same extinction test was administered to all Ss.

Procedure

Feeding schedule and priming. Each S received six drops of the appropriate sucrose concentration (sucrose in tap water) in the food dish in his individual cage for 2 days. All Ss were placed on a 10-g daily feeding of Purina Laboratory Chow for 1 week prior to magazine training. They were fed each day at the time when their training would be completed.

Magazine training. There were 3 days of magazine training, 20 trials per day, with the magazine providing .1 ml of the appropriate sucrose incentive every 30 sec. Separate training cages were used for male and female Ss.

Bar-press training. There were 20 days of bar-press training, with 10 trials administered daily. Training was on a CI 60/60 program. The *controlled interval* (CI) program consisted of the presentation of the bar for a maximum interval of 60 sec, the bar retracting when pressed or automatically at the end of the 60 sec, in either case concluding the trial, and a 60-sec intertrial interval during which the bar remained retracted.

Since the experiment was conducted on Monday through Friday, this phase required 4 weeks. Three Ss were discarded for failure to train, two from Group M-16 and one from Group F-16, leaving a total of 41 Ss in the experiment.

Extinction. Owing to an E error in setting up the automatic controls, usable data from the first day of extinction were not obtained. An additional day of retraining, with 20 trials on the CI 60/60 program, was therefore introduced, and the extinction test administered again on the following day. This test consisted of 20 trials on a CS 30/30 program. The *controlled session* (CS) program consisted of the bar entering for a set period of 30 sec, during which each press operated an empty magazine, and then retracting for an intertrial interval of 30 sec. The latency of the first bar press during each trial and the total number of bar presses per trial were recorded.

RESULTS

Acquisition

Two more or less separate questions were asked of the acquisition data in this experiment. These were: (1) Are there consistent sex differences in either rate to asymptote or asymptotic performance level between the sexes, for either of the two incentive levels? (2) Are there any differences in variability of performance, with the females presumably being more variable owing to the estrous cycle? These questions are considered separately.

Proficiency in learning. Figure 1 shows the mean bar-press latencies for the four groups over the 20 days of training. The

Fig. 1. Mean bar-press latency for the four groups over 20 training days.

statistical analysis of these data confirmed the clearly evident concentration difference ($F = 6.59$, $p < .05$ for 1,36 df), with the 64% groups training more rapidly, but not the very slight trend in favor of the males ($F < 1.00$). Moreover, the two main effects did not interact ($F < 1.00$), nor did sex interact with days ($F < 1.00$), although concentration and days did interact ($F = 4.04$, $p < .01$ for 19,684 df).

Variability in learning. Because each of the groups had reached an asymptotic performance level by the end of Day 12, as shown in Fig. 1, the training data were divided into two sets, preasymptotic (first 12 days) and postasymptotic (last 8 days). Figure 2 shows the mean SD for each of these training phases separately. It is quite apparent that no differences in variability were associated with the 64% incentive. However, the 64% Ss were reliably less variable than the 16% Ss ($F = 5.40$ and 5.96 for first and second phases, respectively, each $p < .05$ for 1,41 df). Although there was a slight trend toward a sex difference in variability for the 16% incentive, with the females apparently more variable early in training and the males more variable later in training, neither of these differences was statistically reliable (neither $F > 1.00$). Moreover, in neither case did the sex and concentration variables interact (neither $F > 1.00$).

Further tests of the variability in learning were made by means

Fig. 2. Mean SDs of latency measures in preasymptotic (first 12 days) training and postasymptotic (last 8 days) training.

of a somewhat simpler measure, the range, computed on both a daily and a weekly basis. The daily measure was simply the difference between the lowest and the highest latency score for the day; the weekly measure was the difference between the lowest and highest daily range during the week. These measures clearly confirmed the difference between the 16% and 64% groups, already shown by the SD measure, and so these groups were considered separately.

The major result of the range analysis for the 64% groups was merely to demonstrate a reliable decrement over both days ($F = 37.17$, $p < .01$ for 19,380 df) and weeks ($F = 8.59$, $p < .01$ for 4,80 df). However, in neither case was the sex variable or the interaction of sex and trials reliable.

For the 16% groups, the daily range analysis showed a similar result, with only the trials factor being reliable ($F = 2.36$, $p < .01$ for 19,65 df). The weekly range analysis, however, showed females to be reliably less variable than males ($F = 5.34$, $p < .05$ for 1,17 df). The weekly means were: for females, 13.59, 18.83, 13.20, 5.71, and 3.60; for males, 24.86, 32.64, 10.57, 11.20, and 9.50. In this analysis neither the trials (weeks) measure nor its interaction with sex was reliable (both F s < 1.00).

In view of the unexpectedly greater variability of male Ss at the 16% sucrose incentive level, it was decided to include one additional variable in the statistical analysis. This was the order of running, that is, whether an S was run each day only after other Ss of the same sex or after some Ss of the other sex. Although separate cages had been used for male and female Ss, it was considered possible that the greater variability shown by males was produced by residual estrual odors within the test box hull affecting males run later on the same day. However, the statistical analysis did not support this proposition. The order variable was not reliable for either the daily or the weekly range measures, for either 16% or 64% groups, nor were any of its interactions with the sex or trials variables reliable.

Extinction

Neither main effect produced reliable differences in extinction, although there was a tendency for the males to make more bar presses (means of 60.3 and 52.8 for the 16% and 64% groups, respectively, compared with 48.1 and 49.1 for the female groups). None of the interactions of either main effect with each other, or with trials, was reliable, except for the interaction between concentration and trials for the bar-press latency measure ($F = 4.73$, $p < .01$ for 3,123 df). This result reflects the somewhat greater tendency for latencies to remain faster for the 16% groups over the latter part of the extinction test.

DISCUSSION

There is no evidence at all in the present data for inferiority in the acquisition of bar pressing by female rats or for their greater variability in bar-press performance. It is clear that the determinants of bar-press proficiency are not sensitive to sex differences and are not seriously affected by the 4- to 5-day estrous cycle.

It may therefore be concluded from this experiment that the sex of the laboratory rats used as Ss in instrumental conditioning studies is not a generally critical variable. While this conclusion must be confined, on the basis of the present data, to the experimental conditions used, particularly to the type of S (hooded rats) and the task (bar pressing), it is very probable that its generality extends well beyond these particular conditions.

Apart from the sex difference in performance tested in the present experiment, choice of one sex or the other will of course depend upon a number of other considerations. On the one hand, the superiority of the females is suggested by their remaining relatively smaller, and therefore somewhat more active, with

increasing age and their presumably less sensitivity to sexual odors, especially those laid down in linear apparatuses. On the other hand, there has been some suggestion in data collected in this laboratory that males are also more sensitive to variations in sucrose concentration (or at least that they are more attracted by higher concentrations). Thus, while the choice of one sex or the other should of course be made in accordance with the particular requirements of the experiment, it is reassuring that use of female rats is not generally contra-indicated for instrumental appetitive learning on the basis of their estrous cycle.

REFERENCES

- COURTNEY, R. J., JR., REID, L. D., & WASDEN, R. E. Suppression of running times by olfactory stimuli. *Psychonomic Science*, 1968, 12, 315-316.
- DAVIS, S. F., & LUDVIGSON, H. W. The "depression effect" and the problem of odor control. *Psychonomic Science*, 1969, 14, 93-94.
- HITT, F. C., GERRALL, A. A., & GIANTRONIO, G. W. Detection of estrous activity cycle by 1-hour samples of running behavior. *Psychonomic Science*, 1968, 10, 159-160.
- LESTER, D. Effects of olfactory stimuli on Y-maze exploration of rats. *Psychonomic Science*, 1968, 12, 97.
- LUDVIGSON, H. W., & SYTSMA, D. The sweet smell of success: Apparent double alternation in the rat. *Psychonomic Science*, 1967, 9, 283-284.
- MARX, M. H., TOMBAUGH, T. N., HATCH, R. S., & TOMBAUGH, J. W. Controlled operant conditioning boxes with discrete-trial programming for multiple experimental use. *Perceptual & Motor Skills*, 1965, 21, 247-254.
- McHOSE, J. H., & LUDVIGSON, H. W. Differential conditioning with nondifferential reinforcement. *Psychonomic Science*, 1966, 6, 485-486.
- MOSS, R. L. Changes in bar-press duration accompanying the estrous cycle. *Journal of Comparative & Physiological Psychology*, 1968, 66, 460-466.

NOTE

1. This research was supported in part by United States Public Health Service Research Grant HD-00895 from the National Institute of Child Health and Human Development, Research Career Award 1-K6-MH-22,023 from the National Institute of Mental Health, and National Science Foundation Research Grant GB 5853.

A measure of pronounciability of CVC trigrams¹

CALVIN F. NODINE² and JAMES V. HARDT, CARNEGIE-MELLON UNIVERSITY, Pittsburgh, Pa. 15213

All 2100 CVC trigrams were scaled for pronounciability by measuring pronunciation latency (P Lat). The resulting distribution of P Lat scores was extremely leptokurtic and positively skewed. Scores ranged from a minimum of .531 sec to a maximum of 1.726 sec. Average P Lat was .81 sec. The relationship of P Lat to Archer meaningfulness was linear; however, the degree of relationship was slight ($r = -.37$). This finding is interpreted as indicating that P Lat is relatively free of bias from such other stimulus attributes as meaningfulness. As such, P Lat is viewed as reflecting a basic processing time for such stimulus materials.

The demonstrated potency of meaningfulness in verbal learning has been amply attested to in several reviews (e.g., Goss & Nodine, 1965; Underwood & Schulz, 1960). One of the important underlying dimensions of meaningfulness is pronounciability. Several studies have reported high positive correlations between rated pronounciability and rated meaningfulness (e.g., Underwood & Schulz, 1960; Wilson & Becknell, 1961). Recently, Gorfain (1967) used pronunciation latency (P Lat) as a measure of pronounciability. Using 144 CVCs from Underwood and Schulz's list, Gorfain obtained an r of $-.66$ between P Lat and rated meaningfulness. This correlation is lower than those reported with rated pronounciability, which led Gorfain to argue that these latter correlations are spuriously high due to significant commonality between ratings of pronounciability and meaningfulness.

One shortcoming of all of these studies, whether they use ratings or P Lat to measure pronounciability, is that they are limited to small subsets of CVC trigrams. This situation may present an unrepresentative picture of the relationship between pronounciability and meaningfulness. The present experiment corrects this situation by scaling the pronounciability of all 2100 CVC trigrams using P Lat as a measure thereof.

METHOD

Materials

All possible CVC trigram combinations were used except those containing duplication of consonants in initial and terminal positions (e.g., "BAB") and those containing "Y" as a vowel (e.g., "BYC"). An alphabetical listing of these 2100 trigrams was randomized by using the ordinal positions of each item as inputs in a computer program that generated new random ordinal positions for each item as outputs. The resulting random listing of CVCs was divided into four sublists of 525 items each.

Subjects

Forty Ss were randomly assigned to one of four sublists containing 525 trigrams each. The order of presentation for each sublist was constant for all Ss. The latter were undergraduates fulfilling course requirements in psychology. There were nine males and one female for each sublist. Each S was run individually and was read the following instructions:

"This is an experiment designed to see how people pronounce nonsense syllables. When the experiment begins, nonsense syllables will appear one at a time on the screen in front of you. I want you to try and pronounce each of these as carefully as possible. Each nonsense syllable will appear on the screen for 1 sec, or until you pronounce it. If you have not pronounced it within 1 sec, it will go off, but you will still have 2 sec to make a response before the next syllable appears. Try to be careful to give the *correct* pronunciation. Do you have any questions?"

After the instructions were read, the sensitivity of the voice key was adjusted to each individual by having S articulate pauses (e.g., ahh, err).

Apparatus and Procedure

Stimulus presentation and data collection were accomplished by means of a DDP-116 computer (Control Data Corporation) and auxilliary equipment that included oscilloscope, voice key, and microphone. The microphone was positioned 6 in. from S's