

A CLINICAL SEARCH FOR LABORATORY REINFORCERS

Beatrice H. Barrett

Working draft, prepared 1979, adapted from:

Analysis and modification of retarded behavior. Paper presented at Massachusetts Psychological Association, Institute for Continuing Studies in Psychology, Boston, October 1968.

Annual Reports of the Behavior Prosthesis Department, 1 July 1970 - 30 June 1971, 1 July 1972 - 30 June 1973, 1 July 1973 - 30 June 1974. Belmont, Mass.: Walter E. Fernald State School, 1971, 1973, 1974.

Reinforcer hierarchies among the moderately and severely retarded. Paper presented at annual meeting of American Academy on Mental Retardation, Toronto, June 1974.

Our principal methodologic goal was to refine a single measurement system that would be 1) applicable throughout the range of retardation and 2) sensitive to habilitatively relevant characteristics of individual behavior patterns regardless of severity of retardation.

For this reason, we constructed our laboratory of eight-inch cinderblock, with urine-impervious floors and metal firedoors. Our apparatus was sturdy enough to withstand assaults with metal chair legs, and our response-defining devices were Lindsley plunger manipulanda sufficiently "overengineered" to be operated reliably over long periods of time by acutely behaviorally disturbed psychotic children and adults.

With these capabilities, we were able to study people with the most disruptive behavior, including those who attacked their environment and those who urinated, defecated and smeared upon it. Figure 1 explains why this was necessary: The prevalence of disruptive behavior (for definition see Barrett, 1971) increases as psychometrically defined retardation increases. Figure 1 also explains why sturdy environments, although necessary, were not sufficient for our task of measuring on a single scale the broadest possible range of retarded behavior. The prevalence of persons showing inadequate response rates in the laboratory also increased with the severity of retardation. Enough severely and profoundly retarded people showed a combination of disruptive and low-rate behavior to warrant modification of our assessment methods.

METHODOLOGIC RATIONALE

In our early studies, most subjects responded at adequate rates for candies or a candy-token mix. But as we included a greater proportion of more severely retarded residents, we encountered subjects who did not work for candy or other edibles. For many of our more severely retarded participants, the edible "goodie" was functionally a curiosity at best and more often simply an object to be done away with or softened and substituted for feces as wall decoration. Most participants who did not consume edibles turned out to be routinely spoon fed (often force-fed) and had never had the opportunity to put food in their own mouths.

To extend our system to include a reasonable sample of severely and profoundly retarded people, we had to determine whether other kinds of events would function as reinforcing consequences. Furthermore, similar studies with chronic psychotic adults had found that low response rates were often attributable to the complexity of the required tasks (Lindsley, 1962). Therefore, to compare potential reinforcers we simplified contingencies by offering opportunities to produce whatever consequence was programmed simply by operating a single plunger manipulandum. Because of the wide range of individual differences shown by severely retarded persons, our studies of reinforcers departed from our systematic plan and became, in effect, a clinical search for rate-sustaining consequences.

To further ensure that the task was within the capabilities of our participants, we modified the Lindsley manipulandum so that it could be programmed to operate by being pushed rather than pulled -- a broader movement topography which did not require grasp-pull-release components. In addition, we provided footswitches that could be tapped or pushed to produce the programmed consequence.

Conjugate scheduling (continuous availability) and episodic scheduling (brief intermittent presentations) were both available. Participants could operate the

apparatus to see five-second presentations of slides (colored wildlife and travel pictures, black and white abstracts, colored abstracts). They could work to hear music (taped rock 'n' roll, classical music, and radio shows) either continuously or in five-second episodes. In a more "social" vein, they could also work to maintain or dim the video or audio channel of TV either with or without the other channel freely (i.e., noncontingently) available. Various combinations of rate-produced increase or decrease of either or both channels enabled us to determine the relative reinforcing power of video and/or audio for a given person (cf. Mira, 1968). In addition we could determine whether the "free" presence of one of the channels potentiated responding for the other. With participants for whom audio and/or video were sustaining consequences, we could also determine whether they could differentiate the plungers when one of them produced something and the other either produced nothing or removed one of the channels.

Participants were 104 residents whose ages ranged from 3 years, 10 months to 43 years, 1 month (mean: 12 years, 4 months). Duration of institutional residence ranged from one day to 33 years (mean: 7 years, 1 month). Thirty-six per cent were classified as borderline to moderately retarded, and 64 per cent as severely or profoundly retarded. Forty-two per cent were considered untestable by conventional methods.

With severely and profoundly retarded participants, we used rate-building procedures if necessary. When events were episodically programmed, we attempted to build to FR 10 (ten responses produced a five-second slide show or five seconds of music). If this worked, we held the schedule constant so that we could compare rates across participants. If FR 10 proved to be a "strain," the ratio was set at whatever produced the person's most stable day-to-day output. For the occasional participant with very high intrasession and intersession rates, the response requirement was raised to FR 20. The same behavior-determined schedule tactic was used when consequences were conjugately programmed. (On a conjugate schedule, music or TV is continuously available if the participant maintains a continuing response rate.) A requirement of 80 responses per minute was effective for most residents, but some strained above 45 responses per minute, while a few easily handled 120 per minute to maintain, say, the loudness of a musical recording or the brightness of a TV picture.

We had planned to give every participant extended opportunities to work for all available consequences, but this proved to be impossible because of transfers, discharges, prolonged illnesses, quarantines, deaths, and other extenuating circumstances, so often encountered in long-term studies. Nevertheless, the majority of our 104 participants were able to show us which consequence maintained their highest output. The sequence of opportunities to work for different reinforcers was determined mainly by scheduling-availability of the various devices; no systematic order was followed. A mixture of candies and tokens or candies and pennies, rather than just candy, was used if and when a participant showed signs of acquiring "exchanging" behavior immediately following each session.

RESULTS

These procedures yielded consequence profiles on a broader range of individual children than we were able to analyze in the more complex two-stimulus, two-response situation (see examples in Figures 2 and 3). We found extensive overlap between the less retarded participants and the severely and profoundly retarded. A number of severely and profoundly retarded participants who emitted very little behavior for candy did maintain high rates of working for colored slides, music,

BOTH
 MEDIAN RATE
 BELOW 100 RESPONSES/HOUR
 AND
 PERSISTENT DISRUPTIVE BEHAVIOR

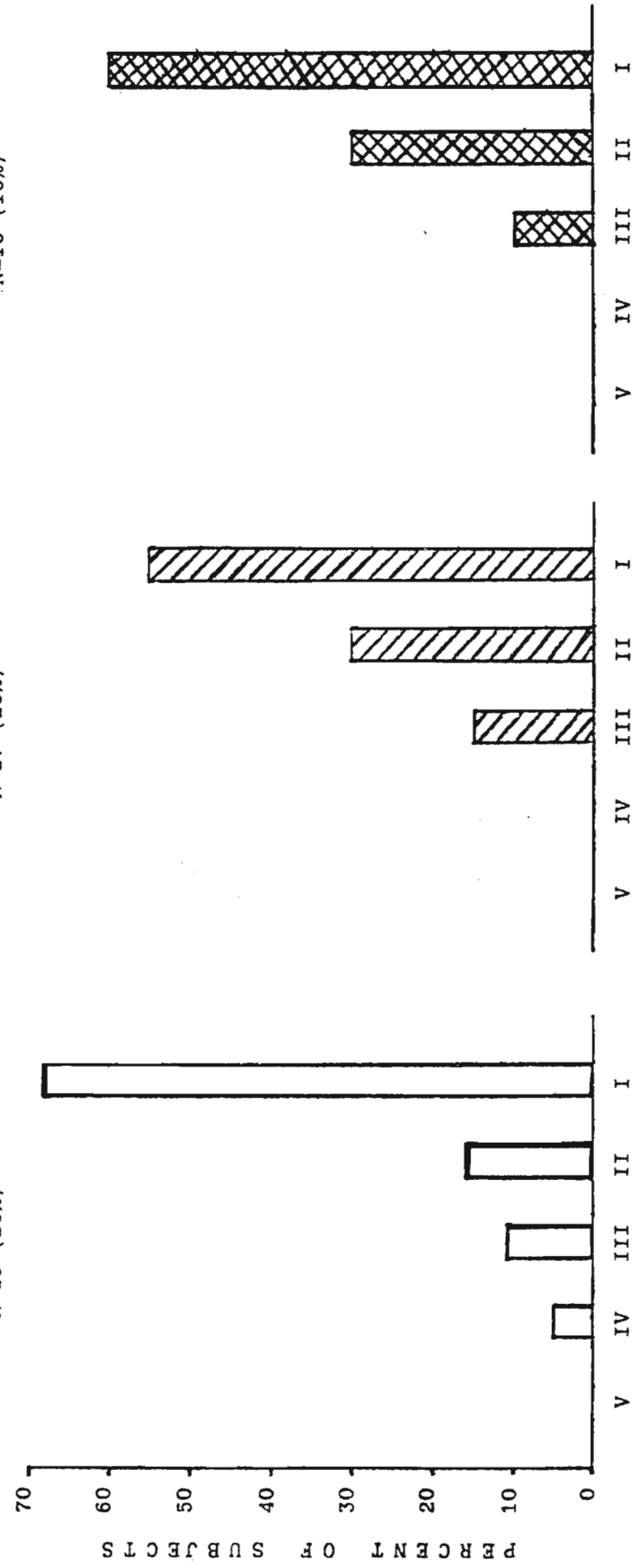
N=10 (10%)

PERSISTENT DISRUPTIVE BEHAVIOR

N=27 (28%)

MEDIAN RATE
 BELOW 100 RESPONSES/HOUR

N=19 (20%)



PSYCHOMETRIC CLASSIFICATION (V-III = borderline to moderate; II-I = severe and profound)

FIGURE 1. DISTRIBUTIONS OF LOW-RATE AND/OR PERSISTENTLY DISRUPTIVE SUBJECTS AMONG PSYCHOMETRIC CATEGORIES.

MIXED DIAGNOSES

MICROCEPHALY

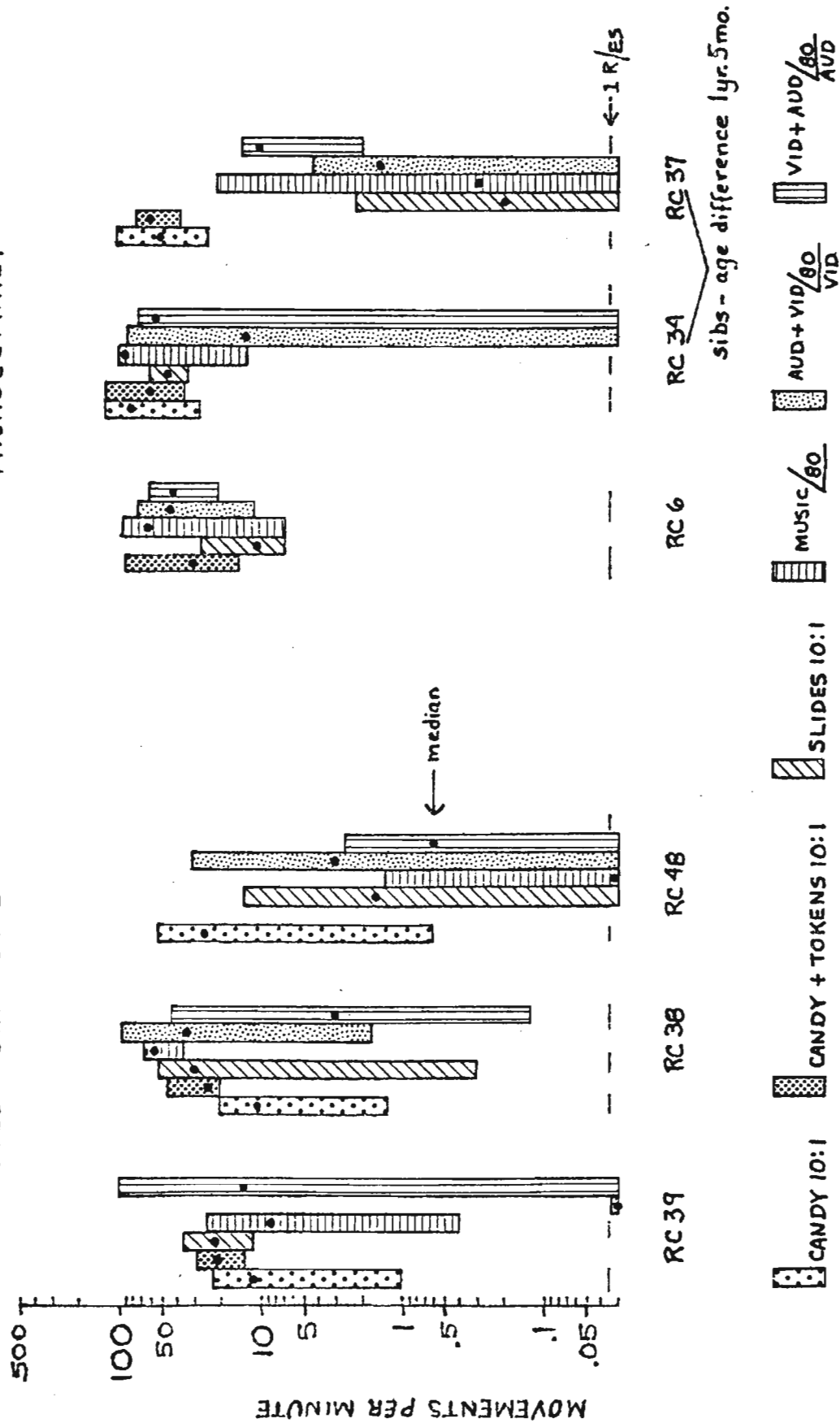


FIGURE 2. INDIVIDUAL PROFILES BASED ON THE RANGES AND MEDIAN RATES OF SIX SUBJECTS WORKING FOR SIX DIFFERENT POTENTIAL REINFORCERS. CANDY, CANDY AND TOKENS, AND COLORED SLIDES WERE AVAILABLE ON AN FR 10 SCHEDULE. MUSIC WAS CONTINUOUSLY AVAILABLE AT OPTIMAL VOLUME IF THE SUBJECT OPERATED THE APPARATUS AT 80 RESPONSES PER MINUTE. THE LAST TWO BARS IN EACH PROFILE REPRESENT RATES OF WORKING FOR THE AUDIO CHANNEL OF COMMERCIAL TV WHILE THE VIDEO CHANNEL WAS FREE, AND VICE VERSA. THE DASHED LINE AT THE BOTTOM OF THE GRAPH REPRESENTS ONE RESPONSE PER 30-MINUTE SESSION.

DOWN'S SYNDROME

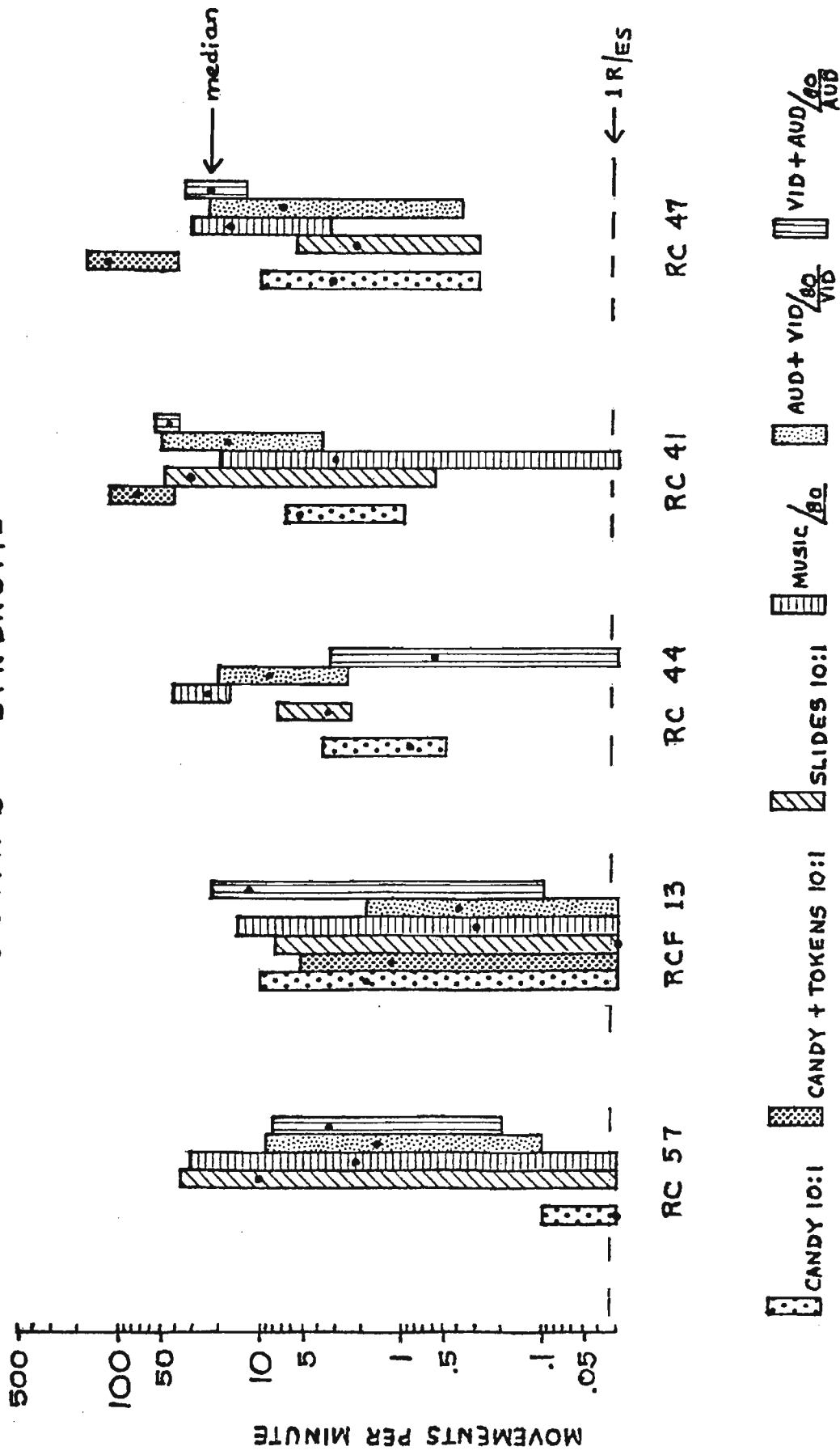


FIGURE 2. INDIVIDUAL PROFILES BASED ON THE RANGES AND MEDIAN RATES OF FIVE SUBJECTS WORKING FOR SIX DIFFERENT POTENTIAL REINFORCERS. (KEY IS EXPLAINED IN LEGEND FOR FIGURE 2.)

or television. Distributions of individuals' median rates of working for their "favorite" reinforcer are given in Table 1. The rank-ordered "popularity" of some of our available consequences is shown in Figure 4.

With this variety of potential reinforcers, we obtained adequate differential rates from 88 (85%) of our total group of 104 subjects, including 57 (82%) of the severely and profoundly retarded children. Fifty-five per cent of our laboratory participants worked at a more sustained rate either for consequences that were not available in our original differentiation-discrimination procedures or for candy when it required only a simple discrimination without any differentiation.

Subgroups Defined by "Favorite" Reinforcer

For each subject we determined the condition that produced his or her highest rate. In the case of discrimination/differentiation studies, only the reinforced rate was considered. The reinforcers were categorized, irrespective of schedule, into 1) candy and pennies or candy and tokens, 2) candy only, 3) auditory events, 4) visual events, and 5) the usual audiovisual combination of television.

The distributions of 88 psychometrically categorized participants in their "favorite" reinforcer category appear in Figure 5. That 73% of the less retarded group worked hardest for a mixture of candy plus either pennies or tokens is not surprising. Tokens were exchangeable for soft drinks, and pennies could be exchanged for coins to operate a candy or drink vending machine or could be saved for later purchases at the institution canteen. Furthermore, return to the dormitory with money often won the admiration of both ward attendants and peers. That the same mixture was attractive to only 26% of the severely and profoundly retarded is consistent with the restrictive ward practices existing at that time.

The most striking characteristic of the distributions in Figure 5 is their marked difference. While a majority of the moderate to borderline residents showed highest response rates for a candy-token or candy-penny mixture, the highest rates of the more severely retarded were nearly evenly distributed across reinforcer categories. It appears that people applying reinforcement technology with the severely and profoundly retarded need a wider variety of potential reinforcers than those working with the less retarded. Fortunately, slide projectors, tape recorders and television sets -- commonly available in homes and classrooms today -- can be used to generate and sustain the behavior of severely and profoundly retarded persons if the devices are controlled by the specific behavior targeted for change.

Psychometrically Associated Behavior Rates

Figure 6 shows how the highest median rates of psychometrically defined subjects are distributed across rate intervals. Both distributions are skewed toward the higher rates. However, in the low-rate intervals (0-39 responses/minute and 40-79 responses/minute), the percentage of severely and profoundly retarded participants markedly exceeds that of the less retarded, while no severely and profoundly retarded participants appear in the highest rate intervals (200 responses/minute and higher). The mid-medians of 88.3/minute for the borderline to moderate group and 25.5/minute for the severe and profound group are significantly different ($p=.03$). The distributions of the two groups over rate intervals are also significantly different ($p=.002$).

Even when each individual was considered in terms of his or her "best" reinforcer, the response rates of severely and profoundly retarded persons, as a group, were significantly lower than those of the moderate to borderline group.

Group Reinforcer Hierarchies

To further summarize our findings on reinforcers, we selected consequences that had been available to comparable numbers of participants in both the moderate-borderline and severe-profound categories. We retained the rank of these consequences from least to most powerful rate-sustainers, as determined by overall group mid-median rates. The question was: Did the two psychometric groups differ with respect to their hierarchies?

Figure 7 shows how the psychometrically categorized participants were distributed with respect to their highest median response rates for consequences that are arranged, from left to right, according to the increasing mid-median rates of the total subject group. The overall, group-determined hierarchy, in order of increasing rate-sustaining power, is:

1. Recorded music, available continuously if the subject maintained a rate of at least 80/minute on a plunger.
2. The audio channel of commercial television, available continuously if the subject maintained his or her most stable plunger rate (\bar{x}), while the video channel was "free" (noncontingent).
3. Five-second episodes of recorded music, produced by every tenth plunger operation.
4. Five-second presentations of colored slides, produced by every tenth plunger operation.
5. The video channel of commercial television, available continuously if the subject maintained his or her most stable plunger rate (\bar{x}), while the audio channel was "free" (noncontingent).
6. A candy, penny or token produced by every tenth plunger operation.

This hierarchy holds up for both of the psychometrically defined groups. A noteworthy finding is that auditory consequences appear to be consistently less powerful rate-sustainers than visual consequences.

Within each consequence category, the mid-median rate of the moderate-borderline group consistently exceeded that of the severe-profound group. Two of the six comparisons are statistically significant with respect to the differences in mid-medians: that for video with free audio ($p < .001$) and that for candy plus pennies or tokens ($p = .04$).

Individual Profiles Showing Marked Selectivity

Group-determined hierarchies may be useful for selecting consequences most likely to sustain behavior in the greatest number of retarded persons. However, a glance at the ranges of individual highest median rates (Figure 7) shows that there are great individual differences.

Table 1
 Rates of Working for "Favorite" Reinforcer
 (All Subjects)

Psychometric Classification	Responses/Minute			
	N	Range	Median	Interquartile Range
Borderline to Moderate	39	0.37 to 300.1	83.2	38.7 to 156.4
Severe and Profound	65	0 to 166.6	15.5	4.2 to 61.7
Both Groups	104	0 to 300.1	41.5	6.9 to 98.8

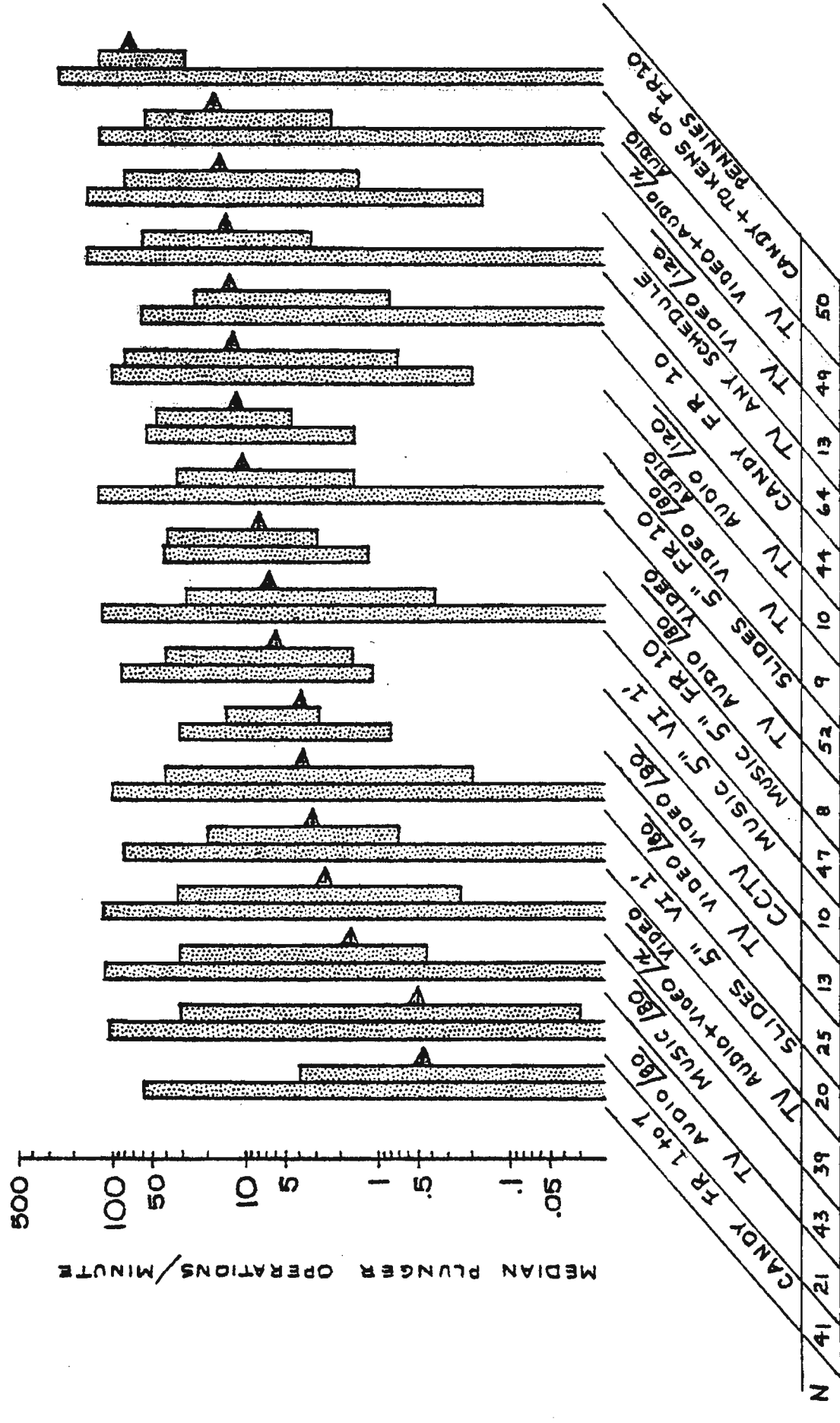


FIGURE 4. CONSEQUENCES RANKED BY GROUP MID-MEDIANS. CCTV IS AN ABBREVIATION FOR TELEVISED SELF-IMAGE OVER A CLOSED-CIRCUIT SYSTEM. CONJUGATE SYMBOLS (/) ARE EXPLAINED IN THE LEGEND FOR FIGURE 2. CONJUGATE RATES LABELED "X" REFER TO WHATEVER PROGRAMMED RATE REQUIREMENT PRODUCED A SUBJECT'S HIGHEST MEDIAN RATE.

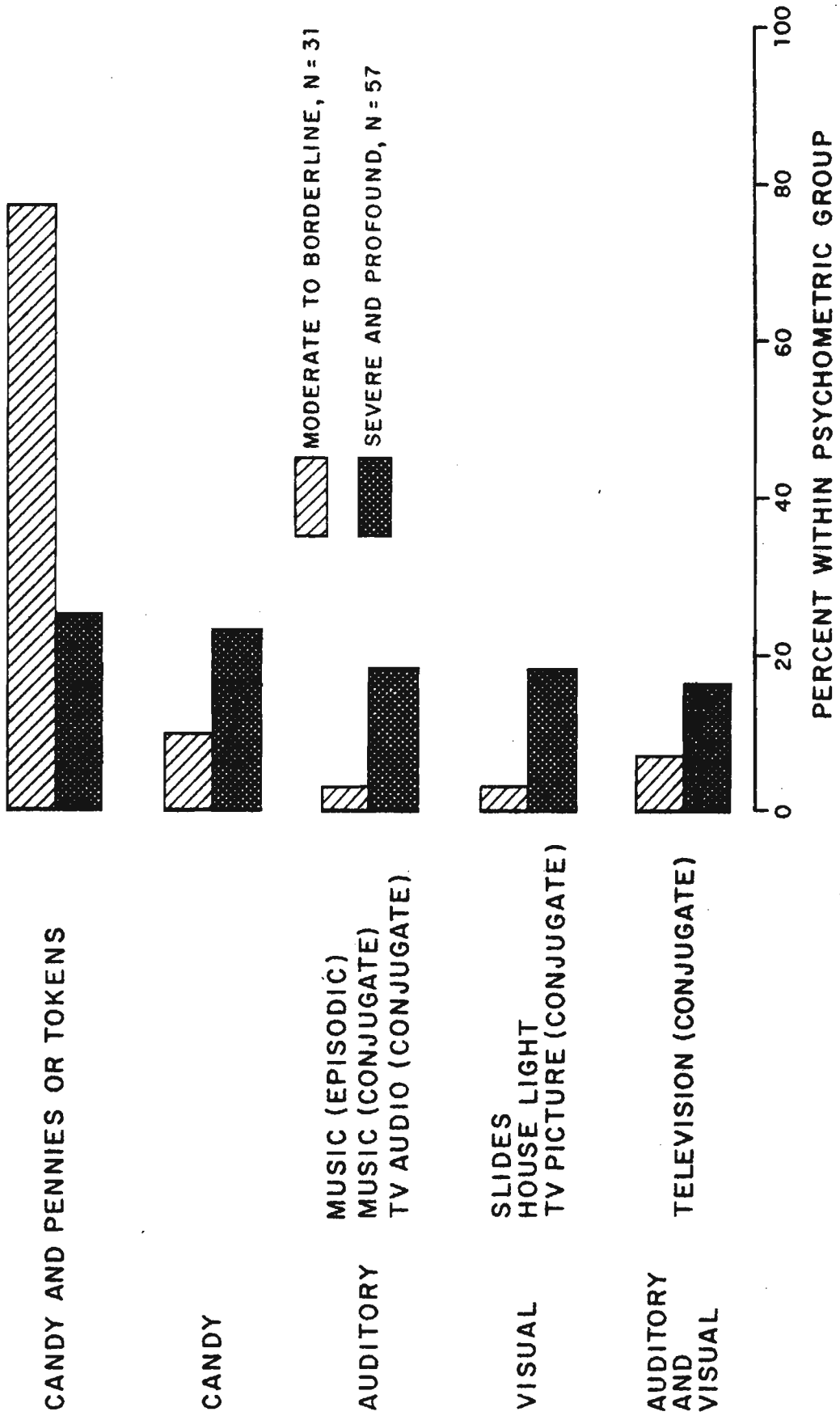


FIGURE 5. DISTRIBUTION OF PSYCHOMETRICALLY CATEGORIZED SUBJECTS IN "FAVORITE" REINFORCER CATEGORIES.

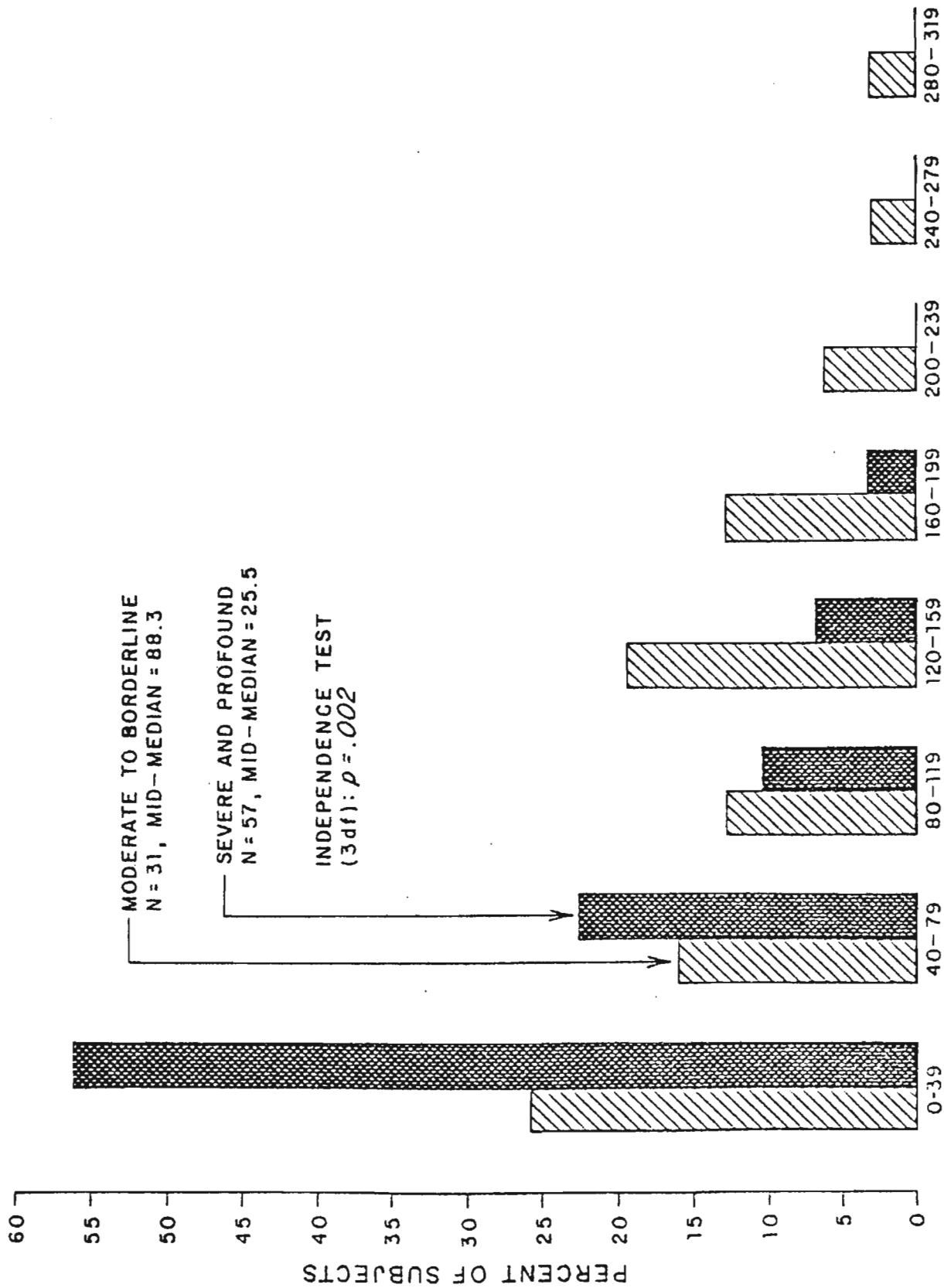


FIGURE 6. DISTRIBUTIONS OF INDIVIDUAL MEDIAN RATES FOR EACH SUBJECT'S "FAVORITE" REINFORCER.

In order to provide potentially useful prognostic information — information that could make training endeavors more individually tailored and therefore more effective, we examined individual consequence profiles of median response rates for indications of selectivity (selective consequence effectiveness). With one exception, we chose only those participants (26 female, 39 male) who had opportunities to work regularly for at least four different types of consequences. Their ages ranged from 46 months to 216 months (mean: 134.0). Age at admission ranged from 0.2 month to 124 months (mean: 46.4). Duration of residency ranged from 0.03 month to 169 months (mean: 87.7).

An individual might show consequence selectivity 1) by working at an unusually high rate for one consequence relative to his or her rates for other consequences and/or 2) by emitting very little behavior to obtain a given consequence relative to his or her rates for other consequences. We grouped individuals with respect to their "peak" and their "weak" rates for given consequences, irrespective of how the consequences were scheduled.

Criteria for "peak" and "weak" rate selectivity were influenced by the marked skewness toward the high end of the "favorite" consequence-sustained rates shown in Figure 7. Because relatively few participants emitted high median response rates, we defined a "peak" rate selective person as one whose highest median rate for a given consequence was at least 1.5 times higher than his or her next highest median rate. Since the highest median rates of the majority of participants are massed toward the low end of the rate distribution, our criterion for a "weak" rate selector was far more stringent. For a person to be categorized as "weak" rate selective, the lowest median rate for a given consequence had to be at least five times lower than the median rate for the next-to-lowest rate sustaining consequence.

Peak-rate selective profiles (N=32) appear in Figures 8, 9, & 10. Seven (10.8% of the 65 participants) were candy-selective; three (4.6%) were music-selective, six (9.2%) emitted "peak" median rates for colored slides, five (7.7%) "peaked" for television, and eleven (16.9%) "peaked" for candy plus tokens or pennies.

These 32 "peak" rate selective people constituted 49.2% of the 65 participants. They included 44.1% of the psychometrically "untestable" participants, 46.7% of the moderate-borderline participants, and 50.0% of the severe-profound group.

Weak-rate selective profiles (N=11) appear in Figure 11. Note that candy plus pennies or tokens did not produce selectively low rates in any of our participants. For one person (RC51), candy was not effective; for one other (RC52), television was not effective. Five persons demonstrated that music was not an effective rate-sustainer for them. Four showed that slides were ineffective in sustaining their behavior. "Weak" rate selective people made up 8.8% of the "untestable" group, 26.7% of the moderate-borderline group, and 14.0% of the severe-profound group.

Selectivity shown by a combination of "peak" and "weak" rates also emerged. Seven of our eleven "weak" rate selective people (63.6%) also appeared as "peak" rate selective. Signified by "P" above their profiles, they constituted 10.8% of the total group of 65. Selectivity, defined by the dual criteria, was not related to psychometric category, to sex, or to psychometric "testability." While the mean age of consequence-selective people (126.3 months) was significantly lower ($p=.02$) than that of the non-selective people (151.3 months), the two groups did not differ significantly with respect to their mean ages at admission or their mean durations of residency.

Selectivity and rate. Of the 65 people who had extended opportunities to work for a variety of consequences, only 10 failed to emit enough behavior to show whether or not they were selective according to our criteria. As shown earlier, the majority of our participants demonstrated relatively low median response rates for their "favorite" (most effective) consequences (see Figure 7). Even though consequences were individualized according to each person's highest median rate, this measure did not adequately dehomogenize individuals in the lower rate ranges.

However, when we examined individual mid-median rates derived from each individual's profile of consequence-sustained median rates, we found that the group whose mid-median rates fell below the combined group mid-median contained a greater proportion of selective than nonselective subjects ($p < .001$). Furthermore, the mid-median rate of the selective group was significantly lower than that of the nonselective group ($.01 < p < .02$).

Thus, in addition to cross-cutting psychometric categories, age at admission, and duration of institutional residency, consequence selectivity derived from repeated opportunities to work for different types of consequences yields clear-cut individual profiles that are especially critical in revealing the individuality of low-rate responders. It is these people who pose the greatest problems in locating long-term-effective reinforcers. By permitting them to show us what consequences sustain their best behavior rates, we have another prescriptively applicable dimension that dehomogenizes this population.

Consequence selectivity and psychometric "testability." Of the 65 participants who had repeated opportunities to earn a variety of consequences, 34 (52.3%) were "untestable" by conventional psychometric methods. Yet, with repeated regular opportunities, 25 (73.5%) of these "untestable" people emitted enough behavior to show distinct consequence profiles. Of the 25, 17 (68%) were selective according to one or both of our criteria: 48% were peak-rate selective; 8% were weak-rate selective; 12% demonstrated both types of selectivity. These strikingly different individual profiles of directly measured behavior furnish yet another indication that, by judiciously programming consequences that are available in most homes and classrooms, we can obtain useful and directly comparable measurements from people whose behavior is not within the range of conventional psychometric methods.

CONCLUDING COMMENTS

Most important for the habilitative relevance of our findings is the fact that all data are based on session-to-session rates under conditions that remained constant for each individual until his or her day-to-day variability ceased to show trends. This approach is in marked contrast to the "try-this-one-day-try-something-else-tomorrow" approach, which fails to permit adequate reinforcer sampling (Ayllon and Azrin, 1968) and which neglects the prolonged adaptation time required when retarded people are exposed to novel stimuli (Barrett and McCormack, 1973).

To check out the clinical relevance of long-term laboratory studies, we conducted an informal study of our laboratory subjects' choice of backup reinforcers in our classroom. From the records of pupil purchases during 175 classroom "store" sessions, we evolved a hierarchy of "backup" reinforcer preferences derived from conditional probabilities of individual choices from among a variety of items offered for "sale" or "rent." Of greatest interest to us were

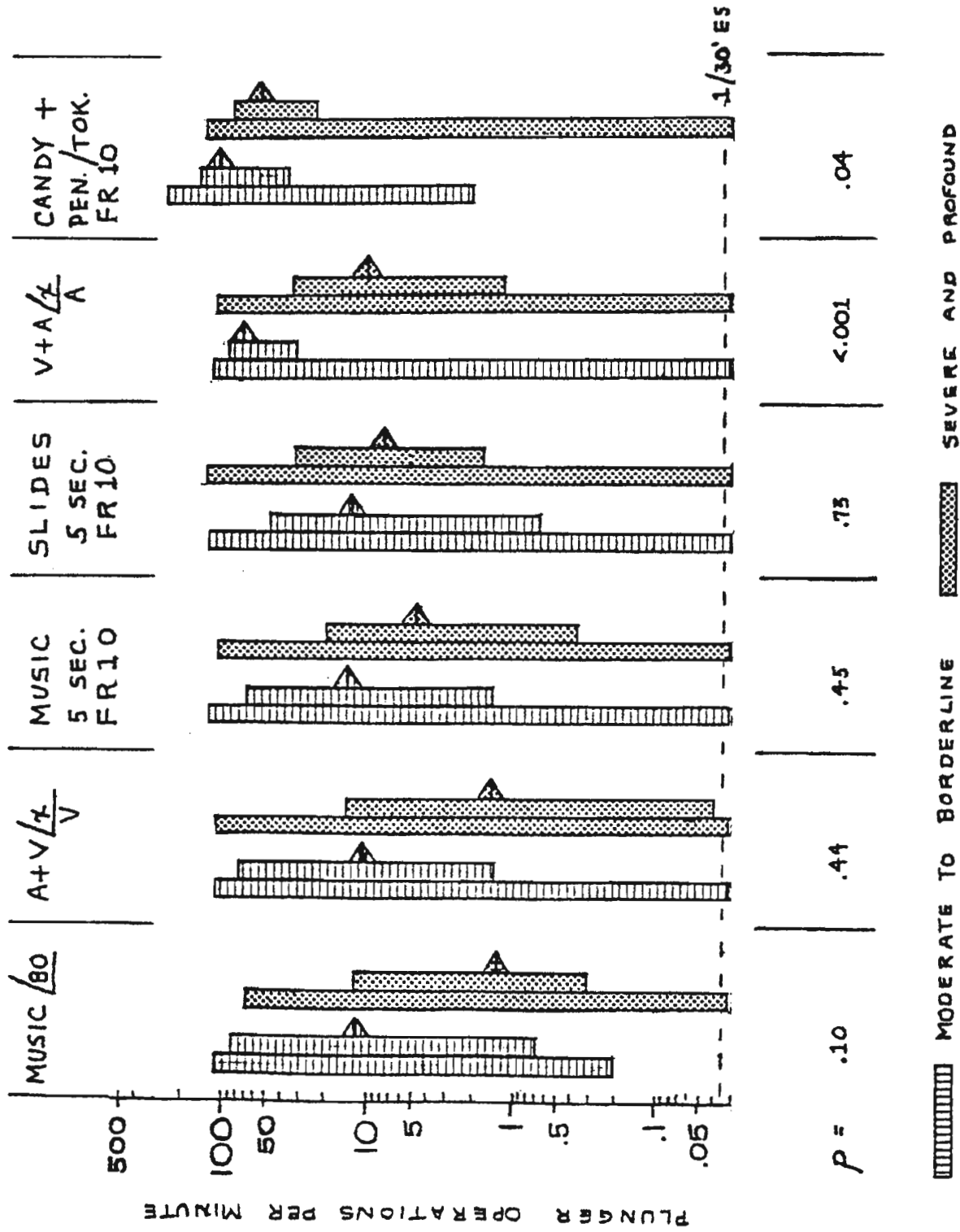


FIGURE 7. DISTRIBUTIONS OF HIGHEST MEDIAN RATES OF PSYCHOMETRICALLY CLASSIFIED SUBJECTS WORKING FOR DIFFERENT CONSEQUENCES (RANKED BY MID-MEDIANS OF COMBINED GROUPS).

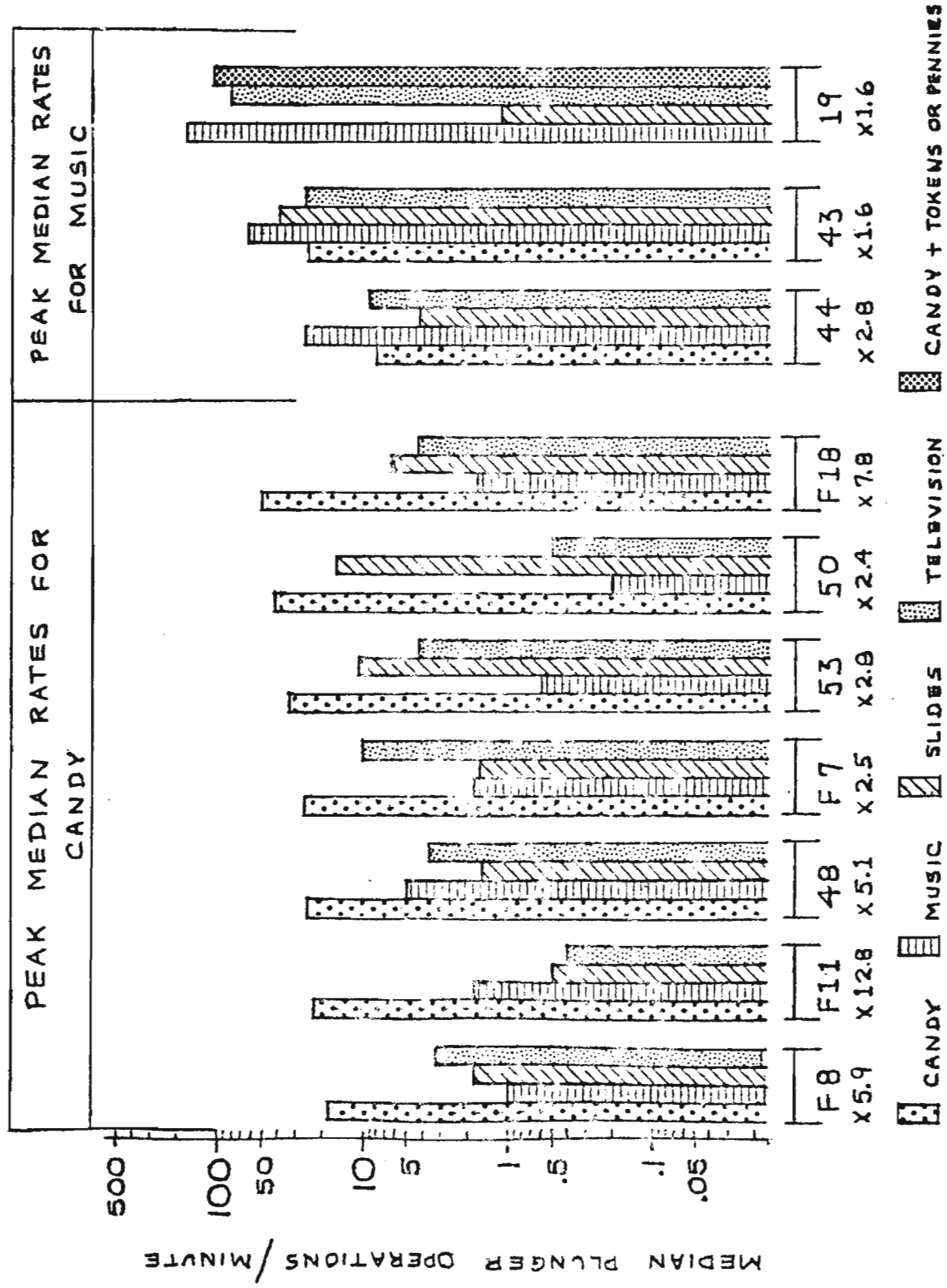


FIGURE 8. CONSEQUENCE EFFECTIVENESS PROFILES OF SELECTIVE PEAK-RATE SUBJECTS.

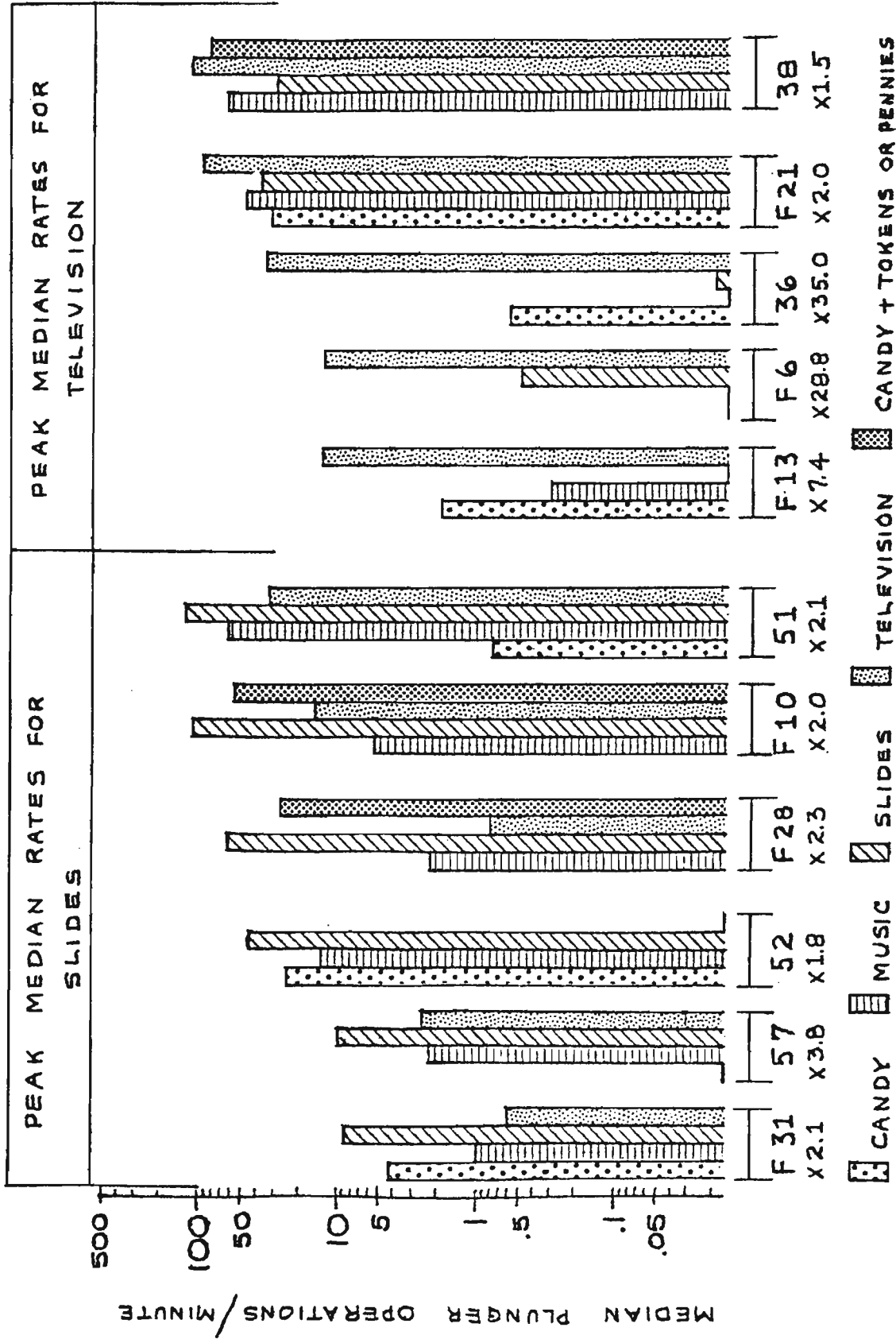


FIGURE 9. CONSEQUENCE EFFECTIVENESS PROFILES OF SELECTIVE PEAK-RATE SUBJECTS.

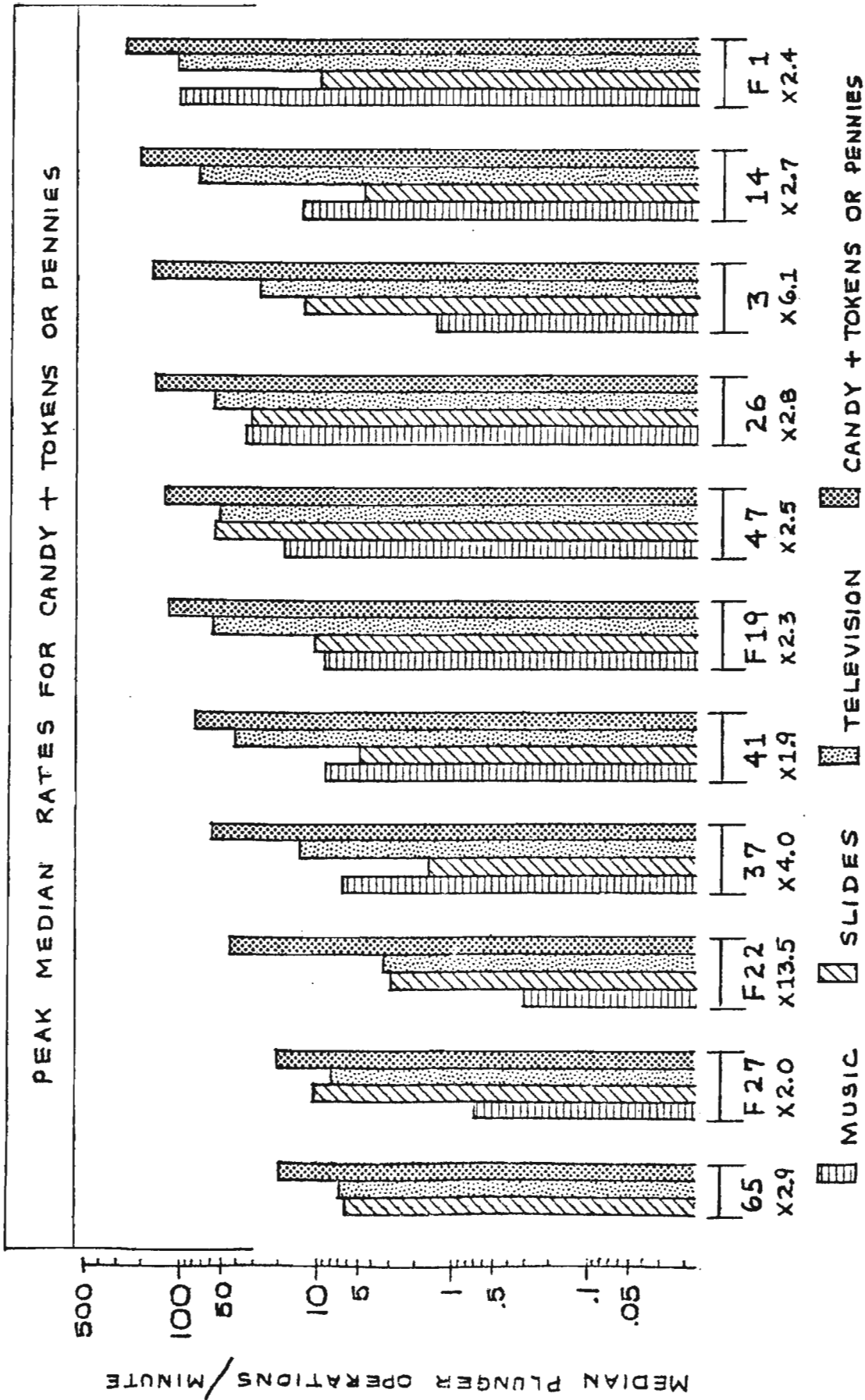


FIGURE 10. CONSEQUENCE EFFECTIVENESS PROFILES OF SELECTIVE PEAK-RATE SUBJECTS.

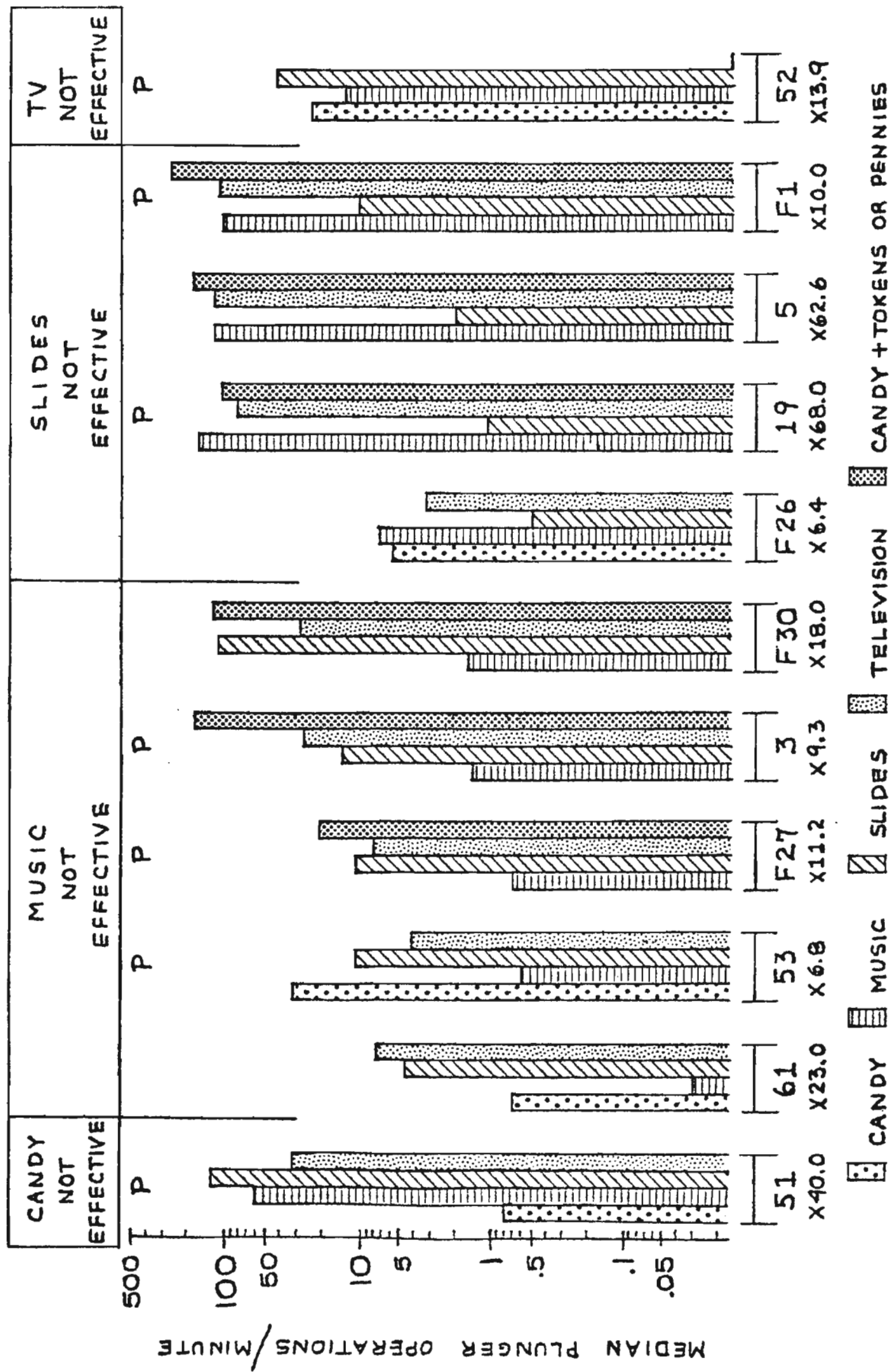


FIGURE 11. CONSEQUENCE EFFECTIVENESS PROFILES OF SELECTIVE WEAK-RATE SUBJECTS.

the items or privileges that corresponded to the reinforcers used in the laboratory -- namely, food and drinks, TV, projected slides, and music, all of which could be purchased or rented at "store" with tokens earned in tutorials. Not counting novelty items occasionally available at the "store," ranks of classroom purchases and laboratory preferences across pupils yielded two identical hierarchies. In the classroom, as well as in the lab, consumable items (food and cold drinks) were most popular. TV was second, projected slides third, and music fourth.

Individual participants' preferences in the two settings were also compared. For 80% of the classroom pupils, their first and/or second most powerful lab reinforcer corresponded to the first and/or second most frequently chosen classroom backup reinforcer. Only 20% of the pupils showed no correspondence in their first and second choices in the two settings.

Obviously, the methods were very different in the two settings; sequential work rates for different consequences and repeated single choices from among many simultaneously available reinforcers are not directly comparable. Given this disparity, we were somewhat surprised to find such consistency of preferences. Does familiarity account for it? Could we have shifted the hierarchy by introducing a broader range of choices? Were individual choices influenced by observing the choices of others? By being in a group setting versus the privacy of a laboratory cubicle? By the intervention of tokens between responses and the act of choosing? Were the individual profiles established by consecutive availability similar to those that might have been established with concurrently available reinforcers in the same setting? Answers to these questions could increase the effectiveness of reinforcement technology in the habilitation of retarded citizens.

REFERENCES

- Ayllon, T., & Azrin, N.H. The token economy: A motivational system for therapy and rehabilitation. New York: Appleton-Century-Crofts, 1968.
- Barrett, B.H. Behavioral differences among an institution's backward residents. Mental Retardation, 1971, 9(1), 4-9.
- Barrett, B.H., & McCormack, J.E., Jr. Varied-teacher tutorials: A tactic for generating credible behavior in severely retarded children. Mental Retardation, 1973, 11(6), 14-19.
- Lindsley, O.R. Operant conditioning methods in diagnosis. In J.H. Nodine & J.H. Moyer (Eds.), Psychosomatic medicine: The first Hahnemann symposium. Philadelphia: Lea & Febiger, 1962. Pp. 41-54.
- Mira, M.P. Individual patterns of looking and listening preferences among learning disabled and normal children. Exceptional Children, 1968, 34, 649-658.