The Psychological Record, 2013, 63, 583–594

GIANT AFRICAN POUCHED RATS (*CRICETOMYS GAMBIANUS*) AS DETECTORS OF TUBERCULOSIS IN HUMAN SPUTUM: TWO OPERATIONAL IMPROVEMENTS

Amanda Mahoney

Department of Psychology, Western Michigan University, and Anti-Persoonsmijnen Ontmijnende Product Ontwikkeling

Timothy L. Edwards

Department of Psychology, Western Michigan University

Bart J. Weetjens, Christophe Cox, Negussie Beyene, and Maureen Jubitana

Anti-Persoonsmijnen Ontmijnende Product Ontwikkeling

Amy Durgin and Alan Poling

Department of Psychology, Western Michigan University, and Anti-Persoonsmijnen Ontmijnende Product Ontwikkeling

Pouched rats can detect Mycobacterium tuberculosis, *which causes tuberculosis, in human sputum. Historically, a phosphate-buffered saline solution was added to sputum in the belief that doing so improved rats' detection of M. tuberculosis, but no relevant data were available. Experiment 1 evaluated rats' performance on samples with and without phosphate-buffered saline solution added. There was no difference in detection accuracy. Adding the solution slows sample processing and will not be done in future operational applications. Experiment 2 compared the performance of rats trained on sputum samples with low versus high concentrations of M. tuberculosis. Training on low-concentration samples improves sensitivity on that sample type. Unfortunately, it is impractical to arrange low-concentration training in the current operational setting, where the rats are used for the second-line screening of samples initially evaluated by microscopy.*

Key words: African pouched rats, tuberculosis, olfactory discrimination, operant conditioning

Although the goal of applied behavior analysts typically is to improve human behavior, altering animal behavior to benefit participating animals or to benefit humans is also a legitimate part of the discipline (Edwards & Poling, 2011). In Tanzania a

The reported research was supported by a grant from UBS Optimus Foundation.

Correspondence concerning this article should be addressed to Alan Poling, PhD, BCBA-D, Department of Psychology, Western Michigan University, Kalamazoo, MI 49008; E-mail: alan.poling@wmich.edu

DOI:10.11133/j.tpr.2013.63.3.012

humanitarian organization called *Anti-Persoonsmijnen Ontmijnende Product Ontwikkeling* (APOPO) uses operant discrimination techniques to train giant African pouched rats to detect landmines and deploys the rats in Mozambique and elsewhere (Poling, Weetjens, Cox, Beyene, et al., 2010). Similar techniques also are used to train the rats to detect *Mycobacterium tuberculosis* by sniffing human sputum (Poling et al., 2011). *Mycobacterium tuberculosis* (MTB) causes tuberculosis (TB), a bacterial disease that typically affects the lungs and is a significant public health concern in resource-poor countries. Recent studies have demonstrated the rats' value for the second-line screening of sputum samples initially evaluated through microscopy (Mahoney et al., 2011; Mahoney et al., 2012; Poling, Weetjens, Cox, Mgode, et al., 2010; Weetjens, Mgode, Davis, Cox, & Beyene, 2009). For example, in 2009 and 2010, the rats screened more than 20,000 patients that had been evaluated by microscopy technicians at Direct Observation of Treatment–Short Course (DOTS) centers, which routinely screen for and treat TB in Tanzania, and increased new case detections by 44% (Poling, Weetjens, Cox, Mgode, et al., 2010) and 42.8% (Mahoney et al., 2011), respectively. Given that TB is a debilitating and often fatal disease and that each person infected with TB typically infects 10 to 15 other people each year (World Health Organization, 2012), these are clinically significant findings.

The results of prior studies indicate that the rats are valuable in second-line screening. A major goal of APOPO's current research is to identify techniques for improving the rats' performance and increasing their operational efficiency, so that sputum samples can be evaluated more quickly and accurately. For example, the rats can evaluate samples very quickly, but the sample processing that is carried out prior to presenting the samples to the rats consumes substantial time, limiting the number of samples available for evaluation. In an effort to shorten the processing time required, a series of experiments was initiated to test various sample presentation methods. A previous study evaluated the ability of the rats to detect the presence of MTB on microscope slides prepared with sputum (Mahoney et al., 2013), a method that would greatly reduce the time required for APOPO's technicians to process samples, because slides are prepared at and are available from DOTS centers. Results indicated that the rats could readily detect MTB at relatively high concentrations, but their specificity was unacceptably low.

The purpose of Experiment 1 was to evaluate another method for reducing sample processing time. Historically, during sample processing at APOPO, a phosphate-buffered saline solution (PBS) was added to each sample and the samples were then heat inactivated to kill MTB and other infectious microorganisms. Some samples contain very little sputum, and APOPO personnel hypothesized that adding PBS would improve MTB detection by the rats, although no relevant data were obtained. APOPO's lab technicians typically prepare 600 to 1,000 sputum samples each Monday, and those samples are subsequently presented to the rats. A time analysis conducted across two consecutive sample processing days indicated that the use of PBS added 70.2 min to the time required to process 600 samples. This time expenditure is warranted only if adding PBS improves detection accuracy. Experiment 1 examined whether it does so.

Experiment 1

Method

Subjects and apparatus. Ten adult pouched rats served as subjects. Alexis, Casey, Kim, Laila, and Queen were from the group evaluated in Experiment 2 (see later section), and Gaitan, Harold, Keane, Mangesho, and Peter were newly selected. The rats were trained and tested in an aluminum and Plexiglas chamber that was 205 cm long, 55 cm wide, and 55 cm high. Ten holes 2.5 cm in diameter were spaced equidistantly apart along the centerline of the chamber floor's long axis. Pots (small plastic cups) containing sputum were placed in a cartridge attached beneath each of the holes. Ethical clearance to conduct the study was obtained from the Tanzanian Ministry of Health.

Sample collection. Two sputum samples were provided by each patient who sought services from one of four DOTS centers in Dar es Salaam, Tanzania. A DOTS center technician analyzed a smear prepared from each sample and recorded whether MTB was present (a DOTS+ sample) or not (a DOTS− sample). For MTB-positive samples, the technician assigned the sample to one of four categories, AFB (a few bacteria), 1+, 2+, or 3+, each indicating progressively larger numbers of bacteria. On average, for all of 2011, 8.52% of the DOTS+ samples were classified as AFB, 31.49% as 1+, 20.18% as 2+, and 39.81% as 3+.

Procedure. Prior to this study, trainers used operant conditioning procedures to teach the rats to indicate samples containing MTB. Training procedures have been described in detail in past publications (Poling et al., 2011; Weetjens, Mgode, Davis, et al., 2009). In brief, habituation, clicker training, and discrimination training were conducted in the first 6 months of the rat's life. During habituation, various people, odors, sounds, and tactile stimuli were presented until the rat did not attempt to escape. Pre-training was conducted in a cage that had three holes placed equidistant in the floor and a food hole measuring 4 cm on the side. During clicker training, a click was followed by access to food (pellets or a banana and pellet mixture). The rats were then trained to approach a hole in the floor of the training cage and stay until the click. Two more holes were opened, and discrimination training began. In this phase, the rats were taught to pause for 5 s above holes containing sputum samples confirmed by microscopy to contain MTB, but not above samples that were not confirmed to contain MTB. The rats were then trained in a 10-hole cage, where the current experiment took place.

Each evaluation session comprised presentation of eight DOTS+ sputum samples, four with PBS and four without PBS. All samples came from the same patients and, for each patient, 5 ml of PBS was added to one of the two samples she or he provided, and no PBS was added to the other sample. The samples to which PBS were added were selected at random. All samples were heat-inactivated at 90°C (Doig, Seagar, Watt, & Forbes, 2002) and then frozen at −20°C until the day of evaluation. In each evaluation session, there were also 62 DOTS− samples. For half of the evaluation sessions, PBS was added to 26 of the 62 DOTS− samples, and for the other half, PBS was added to 36 of the 62 samples. Therefore, overall, half of the DOTS− samples contained PBS. Each rat evaluated 70 samples per day across 4 days. Correct identification responses (pausing for at least 5 s) above a DOTS+ sputum sample were immediately reinforced with a click and food (mashed banana mixed with crushed commercial rat chow) delivered through a plastic tube attached to a syringe. All other responses had no programmed consequences. Identification responses to DOTS+ sputum were correct responses, whereas identification responses to DOTS− sputum samples might have been incorrect responses or correct identifications of MTB missed by microscopy. To evaluate the latter possibility, DOTS− samples that were indicated as positive by two or more rats were analyzed by fluorescent microscopy.

Data analysis. Signal detection theory provides a framework for understanding how properties of the discriminative stimulus (S+), background stimuli, and individual-participant characteristics affect performance on stimulus-discrimination tasks (Green & Swets, 1966). Detection responses in the presence of the S+ are called *hits*, while detection responses in the absence of the S+ are termed *false alarms*. The absence of a response in the presence of the S+ is called a *miss*, and the absence of a response in the absence of the S+ is called a *correct rejection*. A discriminability index called *d'* takes into account both the intensity of the stimulus being detected, as measured by the hit rate, and the level or number of distractor stimuli (referred to as *noise*), as measured by the false alarm rate. The higher the d' value, the easier the stimulus in question is to detect. If adding PBS to human sputum makes MTB easier to detect, then d' and sensitivity (the hit rate, which reflects ability to detect the presence of disease) should be higher when it is added than when it is not added. Moreover, specificity (the correct rejection rate, which reflects ability to detect the absence of disease) should be higher, or not differ significantly, when PBS is added. For all statistical comparisons, effects are considered to be significant if $p < .05$.

Results and Discussion

As shown in Figure 1, adding PBS to the sputum samples did not substantially affect rats' detection accuracy. Mean sensitivity across rats was 75.6% (range 62.5–93.8) with PBS added. Mean sensitivity was also 75.6% (range 68.8–81.3) without PBS. Sensitivity was slightly higher for five rats when PBS was added than when it was not added. The mean difference between sensitivities in the two conditions was not statistically significant (*SE* = 3.85, *t*[9] = .003, *p* = .998). An Anderson-Darling test for normality indicated that the sensitivity scores obtained without PBS were not normally distributed (*AD* = 1.04, *p* = .005). However, a subsequent Wilcoxon signed ranks analysis of the data indicated that the difference between the two medians was not statistically significant (*p* = 1). Mean specificity improved slightly, from 77.9% (range 75–83.9) with PBS to 80.1% (range 74.2–85.5) without PBS, and this effect was evident in 7 of the 10 subjects. The difference in specificity for the two groups was not statistically significant (*SE* = 1.44, *t*[9] = 1.52, *p* = .164), however. The results of the d' analysis are displayed in the bottom frame of Figure 1. The mean d' for samples without PBS was 1.48 (range .99–1.85), while the mean d' for samples with PBS was 1.57 (range 1.15–1.94). The difference between these values, 0.089 (*SE* = .12), is not statistically significant (*t*[9] = −.74, *p* = .469).

Sensitivity and specificity were calculated from 16 DOTS+ samples with PBS and 16 DOTS+ samples without PBS. There were four samples of each type at each of four bacterial counts (AFB, 1+, 2+, and 3+). Table 1 summarizes the rats' evaluations on each bacterial count. Each patient provided two samples; one was treated with PBS and one was not. Most patients received about the same number of rat indications on samples with PBS as they did on samples without PBS. One exception was Patient 1, for whom the sample with PBS was indicated by seven rats, whereas the sample without PBS was not indicated by any rats. There were discrepant results in only one other patient: Patient 6 provided one sample that was indicated by three rats and a second sample that was indicated by eight rats. All other patients received about an equal number of indications on each sample provided.

Table 1
Number of Individual Rats in a Group of 10 That Indicated the Sample to Contain Mycobacterium tuberculosis *in Samples With and Without PBS Added*

Sample type	Patient	With PBS	Without PBS
AFB	1	7	0
	2	10	10
	3	10	10
	4	4	4
1+	5	9	10
	6	3	8
	7	7	8
	8	8	9
2+	9	3	3
	10	10	10
	11	10	10
	12	10	10
3+	13	1	0
	14	9	9
	15	10	10
	16	9	10

This study compared the accuracy of pouched rats as detectors of MTB in sputum samples with and without PBS. As noted, APOPO personnel hypothesized that adding PBS would improve accuracy, and had added PBS to samples in prior investigations, but before the present experiment, they did not systematically test this hypothesis. As Jones (2011a,

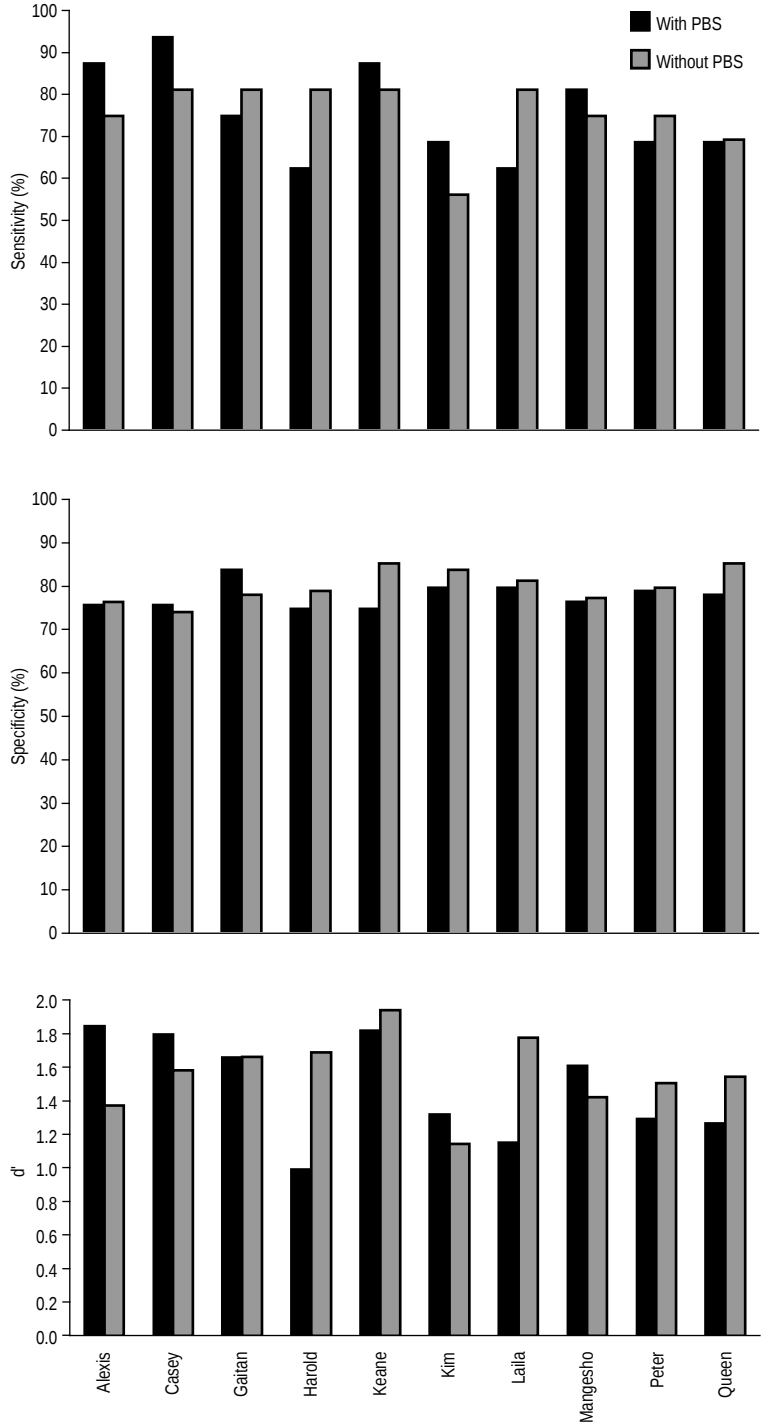

Figure 1. Mean sensitivity, mean specificity, and d' for 10 rats exposed to 16 DOTS+ samples with PBS and 16 DOTS+ samples without PBS.

2011b) pointed out, it is not unusual for people who use animals for humanitarian purposes to incorporate elements of training and testing that are well intentioned but of uncertain

value. As he further indicated, the research strategies and tactics characteristic of behavior analysis provide a powerful and practical tool for evaluating such elements. When that tool was used to test the hypothesis that adding PBS improved rats' performance as MTB detectors, the hypothesis was not confirmed, insofar as results showed no significant difference in detection accuracy as a function of whether or not PBS was added to sputum. Although these "negative" results were obtained in a relatively small study that from a scientific perspective merits replication, they were viewed by APOPO personnel as sufficiently compelling to justify no longer adding PBS to sputum in operational activities. This change has not noticeably affected rats' performance.

Experiment 2

As noted, sputum smear microscopy characteristically has low sensitivity (Steingart et al., 2006). The microscopist's task is easiest when the MTB bacterial count is high and is especially hard when only a few bacilli are present, that is, when the smear should be labeled as AFB or 1+. Therefore, it is especially important that pouched rats used in second-line screening of samples initially evaluated by microscopy consistently detect low-concentration samples. Their ability to do so, however, may be compromised by the kinds of samples used to train them to detect MTB and to sustain their performance. Microscopy tends to miss the lower concentration samples (AFB and 1+) and, because the DOTS+ samples are used for reinforcement samples (i.e., as S+), APOPO's rats have historically been trained largely on the higher concentration samples (2+ and 3+).

Past research indicates that animals will respond to stimuli of a similar intensity to those on which they were trained, while ignoring substantially less intense stimuli. This effect is clearly seen in drug discrimination studies where the training dose determines whether animals generalize to a given lower dose (see Stolerman, Childs, Ford, & Grant, 2011; Young, 1991). It is probable that rats trained primarily to detect high concentrations of MTB do not perform optimally on low-concentration samples. For example, in the Mahoney et al. (2011) study, which involved training largely on higher concentration stimuli, the rats missed six samples that were identified by DOTS microscopy; five of these were 1+ or AFB, whereas only one was the higher 2+ concentration. Training specifically with low-concentration samples should improve the rats' ability to detect such samples, but this possibility has not been evaluated. The purpose of Experiment 2 was to investigate whether training on a preponderance of low-concentration samples improves rats' detection of 1+ samples relative to detection by rats trained in the usual way, that is, with a preponderance of high-concentration samples.

Method

Subjects and apparatus. Ten pouched rats, six females and four males, served as subjects. Five of the rats (Alexis, Casey, Kim, Laila, and Queen) also served as subjects in Experiment 1. Background information on the rats and housing and maintenance procedures are detailed elsewhere (Poling et al., 2011). The experimental sessions were conducted between 0900 hr and 1500 hr. The rat characteristics and experimental chamber were the same as in Experiment 1.

Procedure. The same pre-training procedures and experimental session procedures as described in Experiment 1 were used. Initially, all 10 rats were trained on all bacterial levels (AFB, 1+, 2+, and 3+) for at least 1 year, during which DOTS+ samples were presented as they arrived at APOPO's lab, that is, without regard to bacterial count. Five months before the present evaluation, five of the rats began training on samples classified as AFB or 1+, while the other five rats began training on 1+, 2+, and 3+ samples. The data collected for comparison were obtained using 184 1+ samples, half of which were presented to the lower concentration group and half of which were presented to the higher concentration group across 24 sessions.

Sample collection. Two sputum samples were provided by each patient and were collected from four DOTS centers in Dar es Salaam. As in Experiment 1, a DOTS center technician analyzed a smear prepared from each sample and recorded whether MTB was present (a DOTS+ sample) or not (a DOTS− sample) and, if present, the bacterial count (AFB, 1+, 2+, 3+). Weekly, a lab technician collected the samples and transported them to APOPO's lab for evaluation by the rats.

Sample processing and evaluation. At APOPO, 5 ml of PBS solution was added to each sample and the samples were heat-inactivated at 90°C (Doig et al., 2002) and then frozen at −20°C until the day of evaluation. On Monday of each week, the bacterial count of each sample was entered into a database. Samples were selected from this database for inclusion in the study based solely on the bacterial count and not patient age or gender, sample quality, or other attributes. The database then automatically provided a list for presenting samples so that DOTS+ and DOTS− samples were randomly presented in a given evaluation session.

Each evaluation session involved presenting 70 samples in total; 6 to 10 were DOTS+ samples and the rest DOTS− samples. Samples selected for the low-concentration group had a bacterial count of either AFB or 1+, with 18 exceptions in which 2+ or 3+ were given. These exceptions occurred because there were too few incoming samples with a bacterial count of AFB or 1+. Samples selected for the normal training group had a bacterial count of 1+, 2+, or 3+. In total, the low-concentration group evaluated 70 AFB, 92 1+, 10 2+, 8 3+, and 1,500 negative samples. The normal training group evaluated 0 AFB, 92 1+, 54 2+, 55 3+, and 1,479 negative samples.

Sessions were conducted 3 to 5 days per week. There were no visible markers indicating whether samples were positive or negative. Upon a rat indication, the trainer would inform the data collector, who would state whether the indication was correct. Correct identification responses (pausing for at least 5 s) above a DOTS+ sputum sample were immediately reinforced with a click and food (mashed banana mixed with crushed commercial rat chow) delivered through a plastic tube attached to a syringe. All other responses had no programmed consequences. Identification responses to DOTS+ sputum were correct responses, whereas identification responses to DOTS− sputum samples may have been incorrect responses or correct identifications of MTB missed by microscopy.

Data Analysis. Data for the two groups of rats were compared across the same three measures as in Experiment 1—sensitivity, specificity, and d'—and findings were again considered statistically significant if $p < .05$.

Results and Discussion

Table 2 shows the sensitivity on 1+ samples for all rats prior to low-concentration training. All rats were presented with between 24 and 28 1+ samples within a 1-month period. Mean sensitivity for the five rats that were subsequently exposed to training on AFB and 1+ samples was 64.9% (range 57.7–69.2), whereas mean sensitivity for the rats subsequently trained on 1+, 2+, and 3+ samples was 65.4% (range 56–70.8). The group mean difference was .46 (SE = 3.4), which was not statistically significant ($t[8] = -.13$, $p = .897$). Therefore, there was no between-groups difference prior to the start of this experiment.

The top panel of Figure 2 shows the sensitivity and specificity of all rats on 1+ samples after low-concentration training for one group. All rats were presented with 92 1+ samples across 24 training sessions. Mean sensitivity for the group trained on AFB and 1+ was 82.8% (range 77.8–87.8%), whereas mean sensitivity for the group trained on 1+, 2+, and 3+ samples was 55% (range 44–67.9%). Thus, the between-groups difference in sensitivity was 27.8% (SE = 4.32%), which is statistically significant ($t[8] = 6.44$, $p = .001$). The within-group difference in sensitivity for the rats trained on low-concentration samples prior to and following training was 17.9% (SE = 3.65%), which is also statistically significant ($t[4] = -4.98$, $p = .008$). Mean specificity, shown on the middle panel of the

same figure, was 82% (range 78.1–86.1) for the low-concentration group and 85.6% (range 79.7–89.5%) for the other group. The difference in specificity between the groups, 3.6% (*SE* = 2.13%), was not statistically significant ($t[8] = -1.68$, $p = .134$).

Table 2
Mean Sensitivity on DOTS 1+ Samples for Both Groups During 1 Month of Normal Training

Low concentration				Normal			
Rat	# Hits	# 1+ Samples	Sensitivity	Rat	# Hits	# 1+ Samples	Sensitivity
Alexis	18	28	64.3	Hannah	17	24	70.8
Queen	15	26	57.7	Ray C	17	24	70.8
Laila	19	28	67.9	Onur	14	25	56.0
Casey	17	26	65.4	Rasoul	15	24	62.5
Kim	18	26	69.2	Mary	16	24	66.7
Mean	17.4	26.8	64.9	Mean	15.8	24.2	65.4

The bottom panel of Figure 2 shows the results of the d' analysis, which was calculated using data obtained from 1+ and negative samples during evaluation sessions. The mean d' for the group trained on AFB and 1+ samples was 1.88 (range 1.70–1.97), whereas the mean d' for the group trained on 1+, 2+, and 3+ samples was 1.24 (range 1.12–1.31). The difference between the d's was 0.64 (*SE* = .06), which is statistically significant ($t[8] = 10.79$, $p < .001$).

The goal of this study was to determine whether training on low-concentration samples improved sensitivity on those samples. Results suggest that training on AFB and 1+ samples improved the rats' sensitivity in detecting 1+ samples. After such training, the low-concentration group performed better than they had in baseline and much better than the rats in the other group. Furthermore, the sensitivity of the rats exposed to 1+ or higher samples in the last phase of the study was lower than during baseline, suggesting that training on the full range of sample concentrations (as in baseline) was more effective than training on high-concentration samples only. These findings are unsurprising given that studies with other kinds of discriminative stimuli clearly show that the intensity of training stimuli affects subsequent performance (viz., Stolerman et al., 2011; Young, 1991).

The results of Experiment 2 suggest that training and testing pouched rats to detect MTB under conditions where most opportunities for reinforcement involve presentation of sputum samples with relatively low MTB concentrations improves detection of such samples relative to training under conditions where most opportunities for reinforcement involve presentation of high-concentration samples. A significant limitation of the study, however, is that this effect was demonstrated with 1+ samples, not with both AFB and 1+ samples. This occurred because APOPO receives few AFB samples from the DOTS centers. For example, as noted previously, only 8.52% of the DOTS+ samples received by APOPO were rated as AFB. At the time Experiment 2 was conducted, too few AFB samples were available to train or test rats adequately with such samples. Although it stands to reason that training primarily with 1+ samples, as opposed to 2+ and 3+ samples, would improve detection of AFB samples, confirming this outcome is important because, as noted, it is much harder for microscopists to detect low concentrations of MTB in stained sputum samples than it is to detect high concentrations. Therefore, many of the new case detections by the rats entail their detecting samples with low bacterial counts. This is evident in findings from 2011, where 122 of 508 sputum samples (24%) evaluated as MTB-free at DOTS centers but indicated to be MTB-positive by conventionally trained pouched rats, and confirmed to be so by fluorescent microscopy at APOPO, were AFB samples.

General Discussion

The primary objective in utilizing TB-detection rats for second-line TB screening is to find new TB patients missed by microscopy in a quick and cost-effective fashion. Although a rat can evaluate a sputum sample in less than 10 s, processing samples slows

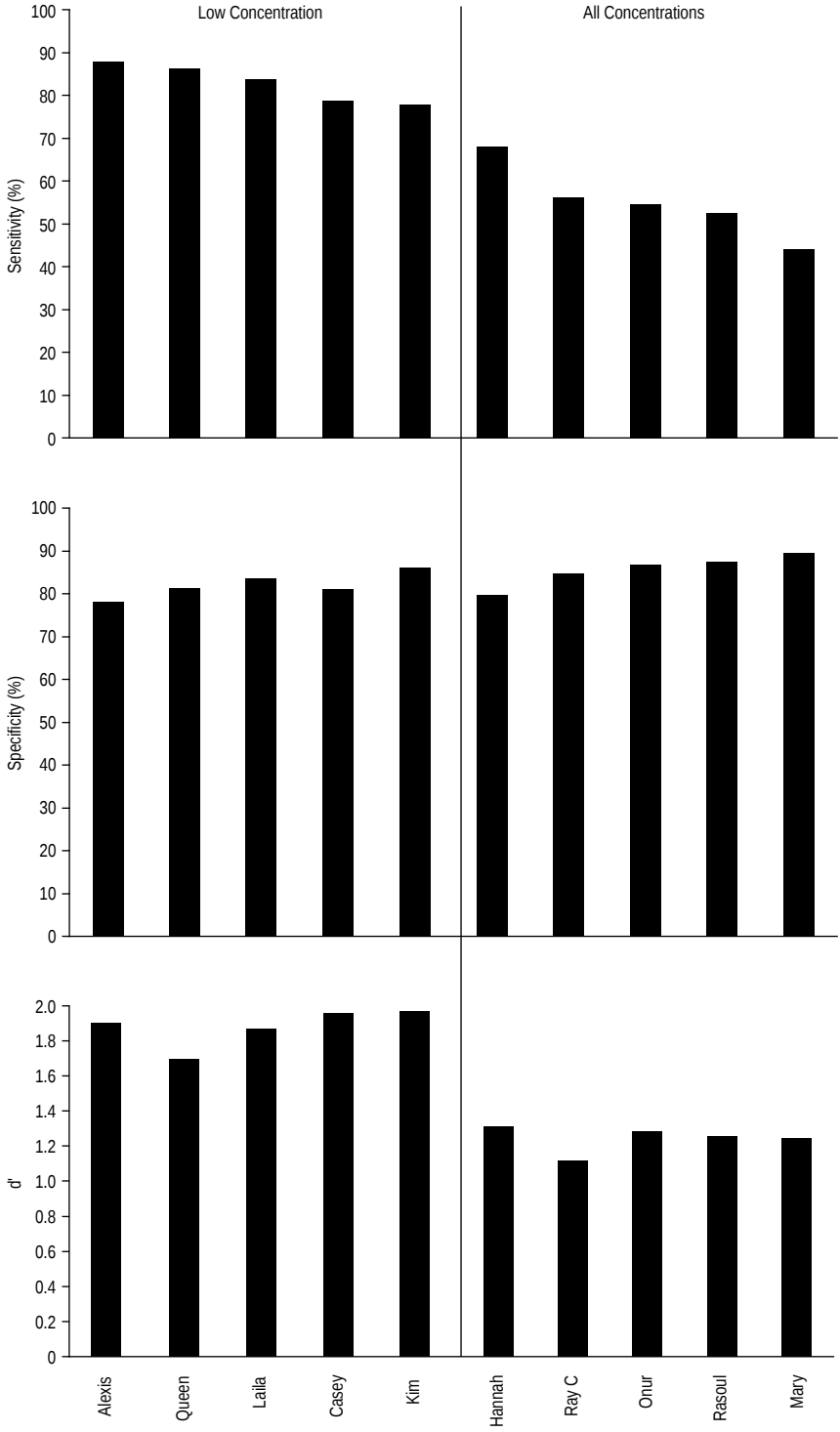

Figure 2. Mean sensitivity, mean specificity, and d' for rats with low-concentration training and rats trained with all concentrations.

the evaluation process. Experiment 1 suggested that one well-intentioned step historically included in processing samples, adding PBS, was of no practical benefit. Therefore, that step is no longer included in APOPO's operational procedures.

Maximizing new case detections requires using the rats in ways that maximize sensitivity while retaining acceptably high specificity. The results of Experiment 2 suggest that training with samples containing relatively low levels of MTB improves sensitivity. Relatively few low-concentration samples typically are available for training, however, and the procedures required for low-concentration training, such as selecting, storing, and presenting low-concentration samples, slow the evaluation process. For these reasons, APOPO has not adopted low-concentration training in its operational activities, although further research evaluating such training is planned.

References

DOIG, C., SEAGAR, A. L., WATT, B., & FORBES, K. J. (2002). The efficacy of the heat killing of *Mycobacterium tuberculosis*. *Journal of Clinical Pathology, 55*, 778–779. doi:10.1136/jcp.55.10.778

EDWARDS, T., & POLING, A. (2011). Animal research in the *Journal of Applied Behavior Analysis*. *Journal of Applied Behavior Analysis, 44*, 409–412. doi:10.1901/jaba.2011.44-409

GREEN, D. M., & SWETS, J. A. (1966). *Signal detection theory and psychophysics*. New York, NY: John Wiley and Sons.

JONES, B. M. (2011a). A sample of behavior-analytic research aimed at developing a system of remotely detecting explosive remnants of war (ERW). In A. Poling & R. Fjellanger (Eds.), *Remote explosives scent tracing* (pp. 81–117). Geneva, Switzerland: GICHD.

JONES, B. M. (2011b). Strategies for the research and development of REST using animals as the primary detectors. In A. Poling & R. Fjellanger (Eds.), *Remote explosives scent tracing* (pp. 62–77). Geneva, Switzerland: GICHD.

MAHONEY, A., WEETJENS, B. J., COX, C., BEYENE, N. W., MGODE, G., JUBITANA, M., … POLING, A. (2011). Using giant African pouched rats to detect tuberculosis in human sputum samples: 2010 findings. *Pan African Medical Journal, 9*.

MAHONEY, A., WEETJENS, B., COX, C., BEYENE, N., REITHER, K., MAKINGI, G., … DURGIN, A. (2012). Pouched rats' detection of tuberculosis in human sputum: Comparison to culture and polymerase chain reaction. *Tuberculosis Research and Treatment*. doi:10.1155/2012/716989

MAHONEY, A., WEETJENS, B. J., COX, C., JUBINITA, M., KAZWALA, R., MFINANGA, G. S., … POLING, A. (2013). Giant African pouched rats as detectors of tuberculosis in human sputum: Comparison of two techniques for sputum presentation. *The Psychological Record, 63*, 21–26. doi:10.1155/2012/716989

POLING, A., WEETJENS, B. J., COX, C., BEYENE, N., BACH, H., & SULLY, A. (2010a). Teaching giant African pouched rats to find landmines: Operant conditioning with real consequences. *Behavior Analysis in Practice, 3*, 19–25.

POLING, A., WEETJENS, B., COX, C., BEYENE, N., DURGIN, A., & MAHONEY, A. (2011). Tuberculosis detection by giant African pouched rats (*Cricetomys gambianus*). *The Behavior Analyst, 34*, 47–54.

POLING, A.,WEETJENS, B., COX, C., MGODE, G., JUBITANA, M., KAZWALA, R., … HUIS IN 'T VELD, D. (2010b). Using giant African rats to detect tuberculosis: 2009 findings. *American Journal of Tropical Medicine and Hygiene, 83*, 1308–1310. doi:10.4269/ajtmh.2010.10-0180

STEINGART, K. R., HENRY, M., NG, V., HOPEWELL, P. C., RAMSAY, A., CUNNINGHAM, J., … PAI, M. (2006). Fluorescence versus conventional sputum smear microscopy for tuberculosis: A systematic review. *Lancet Infectious Diseases, 6*, 570–581. doi:10.1016/S1473-3099(06)70578-3

STOLERMAN, I. P., CHILDS, E., FORD, M. M., & GRANT, K. A. (2011). Role of training dose in drug discrimination: A review. *Behavioural Pharmacology, 22*, 415–429.

WEETJENS, B. J., MGODE, G., DAVIS, B. W., COX, C. L., & BEYENE, N. W. (2009). African giant rats for tuberculosis detection: A novel diagnostic technology. *Social Entrepreneurship, 6*, 39–42.

WORLD HEALTH ORGANIZATION. (2012, October). *Tuberculosis* (Fact Sheet No. 104). Retrieved from http://www.who.int/mediacentre/factsheets/fs104/en/index.html

YOUNG, A. M. (1991). Discriminative stimulus profiles of psychoactive drugs. In N. K. Mello (Ed.), *Substance abuse, behavioral and biological research* (Vol. 4, pp. 139–203). London, England: Jessica Kingsley Publishers.