
APPENDIX VII:

BEHAVIOR EVALUATION

Appendix VII Behavior Evaluation



Increasingly, shelters are using temperament tests to make euthanasia decisions and gather adoption information. Not everyone agrees that this is a good idea.

more



- [What to Pack for Your Pooch](#)
- [Canine Comics](#)
- [Design Your Own Dog Bed](#)
- [The Dinner Party](#)
- Dog is in the Details
- [Summer Travel](#)
- [Word Hounds](#)

In a gathering storm centered on the policies of animal shelters, temperament testing has become a lightning rod. Some resource- and space-starved shelters—which might have once chosen dogs for adoption based on such specious criteria as color, size, age, breed or length of time in the shelter—now use a series of tests that purport to evaluate a dog’s behavior and predict whether the dog will be a good companion for an adopter. Shelters using such tests make several claims for doing so: The dogs they put up for adoption are safer; dogs are selected based on whether they would be good family pets without regard to age or appearance; data gleaned from the tests help shelters find better adoption matches and provide useful information to adopters; and as a result, more people in the community are adopting shelter dogs.

So what’s prompting the firestorm? Several issues. No one advocates putting vicious dogs up for adoption, but many people think good dogs are being declared unadoptable because the tests are unfair and the people administering the tests are not qualified. A common refrain is, “My dog wouldn’t have passed the test.” Further, opponents of temperament testing claim shelters use these tests to hide the truth—that they show low euthanasia rates and high adoption rates by counting only “adoptable” dogs (those that passed the test). This, they believe, deludes a community into believing that there’s no

pet over-population problem, and encourages people to drop off an inconvenient dog at a shelter. Detractors also claim that testing tempts shelters to focus on quick resolution rather than spending in-house resources on prevention and utilizing outside resources such as rescue groups.

Central to all these important and intense issues, though, is the fundamental question: Are temperament tests valid? That is, can testing a dog in a stressful shelter environment predict later behavior of the dog?

Most people advocating tests agree that “temperament” tests, in fact, are not valid because a dog’s “temperament” is subjective. Instead, they prefer calling the tests “behavior evaluations,” because behavior can be seen and described objectively. Two such behavior evaluations, Sue Sternberg’s Assess-a-Pet and Dr. Emily Weiss’ SAFER/Meet Your Match, are the ones most likely to be used by shelters because information about these tests is readily available through workshops, seminars, books, and videos as well as from such organizations as the American Humane Association and the American Society for the Prevention of Cruelty to Animals (ASPCA).

Assess-a-Pet

Assess-a-Pet, a step-by-step behavior evaluation that takes about 15 minutes, was developed by Sue Sternberg. Sternberg based the test on her 23 years of dog behavior experience, and has refined it over the past 11 years at the nonprofit shelter she founded in upstate New York, Rondout Valley Animals for Adoption.

“The purpose of the test is to find the gems that don’t often come in gemlike packages,” Sternberg says. “I wanted to develop a test that would reveal what the dog would be like with the average adopter, not with a professional dog trainer.” It begins with hands-off observation in which the tester looks for sociable or nonsociable responses, and progresses to evaluations for play, arousal, resource guarding, behavior with cats and mental sensitivity. The test uses the infamous Assess-a-Hand, an artificial hand on a stick that allows someone testing for resource guarding to safely approach, pet and then try to pull a food dish or chew toy away from a dog. Among other recommendations, Sternberg advises shelters to wait two to four days before testing and have two trained people perform the test.

Assess-a-Pet is not a simple pass/fail test; in most parts of the evaluation, the tester selects among a range of responses and also adds observations. For example, the four responses to a test during which the tester strokes the back of the dog are: moves toward tester in at least two out of three strokes, stays in same spot, moves away from tester, or freezes and becomes more aroused. Although some dogs have extreme responses, most responses land in a gray area.

“Mostly, the tests give us information that helps us determine who we can put the dog with,” says Trish King, director of behavior and training at the Marin Humane Society (in northern California), which bases their behavior evaluations on the Assess-a-Pet test. “If a dog is problematic in one area but fantastic in others, we will go out of our way to place that dog because we have the room and the training facility. Unfortunately, other places don’t.” At the Marin Humane Society, virtually all dogs are held for three to four days

before any testing, walked outside in a lawn area to relieve themselves first and tested in a quiet room away from the kennels by two people (one of whom has gone through a full apprenticeship program). Any dog that fails—about 5 percent according to King—is retested at least once within three days, and all dogs who show health problems are tested again once they're healthy.

SAFER/Meet Your Match

Emily Weiss, PhD, divides behavior evaluation into two parts, the SAFER (Safety Assessment for Evaluating Rehoming) test, and the Meet Your Match program, both developed at the request of the Kansas Humane Society. SAFER, a six-part test designed to evaluate aggression quickly (in about six minutes), also uses Sternberg's Assess-a-Hand for food guarding. In this evaluation, a dog is given an A, B, C, D or F in each part. For example, during the sensitivity test, in which the handler kneads and squeezes large handfuls of skin from the dog's ears to its tail, if the dog accepts the touch, it gets an A; if it quickly turns toward the handler's hand and mouths with little to moderate pressure, a C; if it growls or tries to bite, an F. Weiss recommends that all the tests be conducted by two people and video-taped. As with Sternberg's test, each shelter determines, based on its resources, what combination of grades determines adopt-ability. After a dog is SAFER tested, the shelter might then use Weiss's Meet Your Match program to evaluate the needs of individual dogs and gather information from potential adopters to find compatible homes.

The ASPCA in New York, which receives dogs from their humane law enforcement officers, from the NYC Animal Care & Control, and from owner surrenders, uses the SAFER test to determine whether to accept owner-surrendered dogs. "The ACC dogs that we take have already been evaluated," says Pamela Reid, PhD, director of the Animal Behavior Center. "But for the owner surrenders, we use the SAFER test to get a quick assessment. We've raised the bar on which of these dogs we're willing to accept because we already get a lot of problem dogs from humane law enforcement." Once a dog has been in the shelter a few days, it's given a full evaluation using parts of a 140-test-item behavior evaluation developed by Dr. Amy Marder, a veterinarian now with the Animal Rescue League of Boston. "The full test took an hour-and-a-half," says Reid. "So, we're using a pared-down version based on her research that includes only the parts that are predictive of behavior in the home."

San Francisco SPCA

The San Francisco SPCA began developing its own behavior evaluation test when Jean Donaldson, PhD, joined the shelter in 1999. "Sue [Sternberg] is a pioneer, and using her test is a better way of choosing dogs than deciding to keep the ones that have been in the shelter the longest or shortest time," she says, "but we need tests that are scientifically proven to be reliable and valid. We couldn't get Sue's test past the reliability issue, and four of her five unadoptable dogs did fine. We adopted out three and did behavior modification on one."

So, the SF/SPCA devised its own test. "We sat down with all our trainers, decided what we were going to accept or not going to accept, defined our terms, and created a test with

objective scoring,” Donaldson says. “We’ve got to have an objective test or our data becomes junk.”

Instead of asking if a dog is friendly, for example, they ask if the dog approached a handler within X number of seconds; if it growled for three seconds when a stimulus was within six feet on the right side; and, as the stimulus came closer, did the dog snap or continue to growl. “We’re checking boxes and at the end we can see if the dog is above or below our criteria for an adoptable dog,” says Donaldson, who notes that dogs often pass the test with suggestions for behavior modification. “Because the criteria were agreed upon by all people in the shelter, and the result is the same whether I test, you test, the test happens this week or next week, no one is forced into a good position.”

To determine reliability, they tested their method in two ways: The dog was retested (without behavior modification) a week later by the original tester and the results were compared; and three to five testers tested the dog independently and those results were compared. Because results were the same, the test was deemed reliable.

As for valid? “We keep records on all the dogs, but what has to happen and has not happened is the follow-up,” Donaldson says. “The issue with our test and with all the evaluations is that we haven’t crunched enough follow-up numbers. We have to say we really don’t know.”

Some data on temperament tests is slowly becoming available, though.

Testing the Tests

Weiss, for example, followed two groups of dogs at the Kansas Humane Society through adoption or euthanasia. One group was given the SAFER test; the other given health checks but not a behavior evaluation. Of the 141 dogs, 12 were euthanized for behavior reasons and of those, only four were in the SAFER tested group. A follow-up phone survey three weeks after the dogs were adopted determined that 36 dogs from the untested group showed aggression compared to eight from the SAFER-tested group. “We repeated the test about six months later and got similar results,” says Weiss. “After that, they were not comfortable putting dogs up for adoption that hadn’t been tested.”

She has also begun evaluating dogs in boarding kennels to see whether the tests are as valid for dogs with homes as for dogs in shelters. “On dogs already in loving homes, SAFER is proving to be predictive of aggression and nonaggression,” she says. “While we are still collecting and analyzing the data, early reports indicate a strong predictability.”

In a separate study, Dr. Marder has been looking at the results of follow-up phone surveys for 70 adopted dogs that were assessed at the ASPCA using her 140-test-item behavioral evaluation. “I was seeing dogs put to sleep that were like dogs in my private practice,” she says. “The owners were working on the problems and the dogs were doing fine. So, I wanted to find out which tests in the behavioral evaluation were predictive of behaviors in the home.”

Each test-item in the evaluation called for objective observations: Evaluators described the placement of a dog’s ears, for example, rather than classifying a dog as “happy.” And,

the evaluation as a whole was tested and determined to be reliable: results were the same regardless of who did the testing.

To organize the study, Dr. Marder grouped the test items into such categories as possessive behavior, handling, protective behavior, cage behavior and response to fearful stimuli. The dogs' responses were also categorized by such behavior as aggressive, friendly and fearful. The phone surveys made one, two, three and six months after adoption asked about these categories.

In "Pick of the Shelter," (Bark, Fall '03) Patricia McConnell, PhD, wrote, "It is impossible to perfectly predict the behavior of a dog in one context when you're doing the evaluation in another. Period. End of sentence. Impossible." Dr. Marder's results show that this statement is true.

Rather than trying to draw a perfect correlation between a shelter test and behavior in the home, Dr. Marder decided to look at how well (how perfectly) a test predicted behavior, in the same way, for example, that results of an SAT test predict academic success or failure.

Once her numbers were crunched, she concluded that none of the individual test items were 100 percent predictive; each test only indicated tendencies. She also determined that the ability of any test to predict behavior changed over time. "The dogs change in two directions, an increase in behavior or decrease in behavior," she says, and recommends that other information, such as intake profiles and the behavior of the dog in the shelter, also guide predictions and triage decisions.

With this in mind and looking at the broad picture, Dr. Marder's analysis shows that if a dog growled, snapped or bit during any test in the shelter evaluation, the dog was more likely than not to exhibit one of these behaviors again after adoption. But, importantly, by digging deeper into the numbers, she saw that growling during any test at the shelter did not predict snapping or biting after adoption.

When considering categories of behavior, she found three for which positive tests were moderately predictive: possessive aggression, protective behavior and mouthing. That is, if a dog lifted a lip, growled, snapped or bit over food, rawhide or a bed during the test, the dog was likely to show some form of possessive aggression after adoption. Similarly, dogs who lifted a lip, growled, barked, snapped or bit when approached or threatened by a stranger (protective behavior) were likely to show territorial behavior after adoption. And dogs that mouthed during the test were likely to mouth after adoption.

Somewhat predictive were positive responses in categories having to do with aggression to children (dogs were tested with a toddler doll), interdog aggression and separation anxiety. And if a dog showed cage aggression in the shelter, it was somewhat likely to exhibit territorial behavior after adoption.

Of course, what the dog doesn't do during an evaluation is also important. For example, dogs who did not show possessive aggression, separation anxiety or fear of people during the test were not likely to have these behaviors pop up after adoption, either. And a dog's

friendliness, or lack thereof, in the shelter tended to be the same after adoption. The number crunching continues as she readies the data for publication.

Testing Assess-a-Pet

In addition to Weiss and Marder, two researchers who have been compiling data for behavior assessments based on Sue Sternberg's test—Janet Smith at the Capital Area Humane Society in Lansing, Michigan, and Kelley Bollen, a behaviorist with the Massachusetts SPCA—are about to release their findings.

For her first study, Smith tracked 839 behaviorally assessed dogs adopted over a two-year period. The results, which she's planning to present at the HSUS/Animal Care Expo in March, show that dogs put into a level-one category (no restrictions) after the behavior assessment stayed in the shelter an average of six days, level-two dogs (restrictions such as homes with older children) stayed an average of nine days, and level-three dogs (more difficult issues) stayed 14 days. Some of the level-one dogs were returned and adopted out again, but none were euthanized. On the other hand, 3 percent of the level-two dogs and 7 percent of the level-three dogs were returned and euthanized (or euthanized elsewhere) for behavior problems. "Our return rate has decreased since implementing an assessment process," she says. "We are making better matches and our euthanasia rate has not increased." Smith believes that because of temperament testing, the shelter is putting safer dogs up for adoption.

Bollen tracked 2,017 dogs that she tested personally with Assess-a-Pet using follow-up calls at six months for every dog and at one year for random dogs. "I tried to do as many components of the test as I could, whether or not the dog was aggressive during the test," she says. Bollen, who hopes to have her results published in a peer-reviewed journal, was unwilling to release actual statistics at this time, but did share some general results.

"I found that if a dog showed overt aggression that caused it to fail one part of the test, it was likely to show overt aggression in other parts of the test," she says. And, of the dogs she deemed adoptable, a high majority showed no aggression after adoption. "My results show that the temperament test does identify dogs that have a tendency to exhibit aggression in certain situations. Performing the test reduces returns because we reduce the number of aggressive dogs who are placed back into the community, and it allows us to make better placements. And, lastly, borderline dogs, the ones that showed behaviors of concern during the temperament test but were adopted out, were more likely to exhibit behavior problems or aggression post-adoption."

The results sound encouraging; however, canine behaviorist Dr. Karen Overall, who is on the faculty of the University of Pennsylvania's School of Medicine, casts a skeptical eye on temperament testing and the data being presented. "I think Amy Marder's work has a lot of potential because she's asking about probability, about how consistent the dog's behavior is over time," she says. "I'm a scientist. Before I can look at findings, the test has to be repeatable and reliable and there has to be objective criteria. We have to codify the behavior ... where the dog's ears are, if there's vocalization, and if so, whether it starts low and goes up or goes down, where the feet are, what the hair is doing. And context matters. The people who use the Assess-a-Hand do so to have a safe way to reach toward the animal, but the first set of conditions is whether your test instrument is valid.

This test object doesn't mirror the real world, so the answer has to be no. So, don't tell me a dog growled.

"I'm not saying there aren't factors in these tests that will be predictive, but they may not predict what people think," Overall adds. "When I review the tests, I see spurious correlations."

Dr. Overall isn't alone among behaviorists in questioning the tests. "We do our damndest to find appropriate placements," says Reid. "The test gives us just one snapshot of behavior. We've had dogs that aren't good on the evaluation but were fine with the people who were walking them and cleaning the cages. So we take that into consideration."

Reid joins her colleagues in calling for more research. "The two things that are missing are, first, more studies and greater numbers," she says. "And second, we need information about dogs that fail an evaluation in some way, undergo rehabilitation and get adopted out. We need to know whether the behaviors resurface."

Adds Donaldson, "The anti-testing people are so incredibly well-meaning. I know where they're coming from. You run a test, adopt the dog anyway, and the dog is fine. Clearly there are problems with the tests, but it could be that some tests are valid, that some parts of the tests may have good predictive value. The preliminary results from tests by Emily [Weiss] and Amy [Marder] have value and are a tantalizing reinforcement for some things, but we have to get funding for more research. Before we can save all the dogs, we have to triage; we have to save the maximum number of dogs in a way that makes sense. If testing is not the way, if it turns out that there is no way to test that's adequately valid, then we'll need to stop banging our heads on the testing wall. But then what will we go on?"

Implicit in the work these researchers and behaviorists are doing and in the worries people inside and outside the shelter system have about temperament testing is their concern for the community and for the dogs. Pete Miller, a shelter supervisor at Santa Barbara County Animal Services and a 20-year veteran of the shelter system who believes temperament tests are a necessary part of good sheltering practice, perhaps puts this best: "When a dog dies in an animal shelter, it almost doesn't matter whether the dog was an old favorite or a hopeless case of a violent animal that never had a chance; the dog was alive one second, and literally gone the next. Everything it ever was and every possibility for what it would have been and done—gone in a second. It's the actual fact of the real loss and what it means to kill that needs to weigh most and is the reason there should never be a formula that tries to remove the responsibility from a person or dim the reality of what it means to take away a life.