Mouse Taste Preference Tests: Why Only Two Bottles?

Michael G. Tordoff and **Alexander A. Bachmanov**
Monell Chemical Senses Center, 3500 Market Street, Philadelphia, PA 19104-3308, USA

Abstract

Two-bottle tests have been used extensively to measure the preference for taste and nutrient solutions but there has been little work with tests involving more than two bottles. Here, we compare the results obtained in two-bottle tests with those obtained in three- and six-bottle tests. In Experiment 1, we measured the preferences for 2 mM saccharin, 50 mM citric acid, 0.3 mM quinine hydrochloride and 75 mM NaCl displayed by 129X1/SvJ (129) and C57BL/6J (B6) mice. Mice drank more taste solution when they received two bottles providing taste solution and one providing water than when they received either a standard two-bottle test or two bottles providing water and one providing taste solution. The three-bottle tests also revealed the left spout side preferences of the 129 strain and were generally better at distinguishing between the 129 and B6 strains (i.e. were more sensitive) than were the two-bottle tests. In Experiment 2, we measured intakes and preferences in tests with six bottles, with one, two, three, four or five containing 75 mM NaCl and the rest containing water. NaCl preferences were monotonically related to the number of NaCl spouts available. A follow-up experiment found similar results whether the index of ingestion was volume intakes or licks. This argues that spillage cannot account for the effect of spout number on taste solution intake. Together, the results suggest that (i) the number of bottles of taste solution and water has a profound influence on taste solution intake and preference, and (ii) three-bottle tests may be more sensitive than two-bottle tests in many circumstances.

Keywords

129X1/SvJ; C57BL/6J; citric acid; NaCl; quinine hydrochloride; saccharin; side preferences; spillage; taste tests

Introduction

Many investigators have used two-bottle choice tests to investigate rodent preferences for nutrients and taste solutions. Although the large majority of such studies involve rats, recently, there has been a strong impetus to use mice as subjects because of their advantages for the genetic dissection of complex behaviors. The use of mice presents several challenges. First, they drink less than do rats. This is a challenge because it is difficult to measure small volumes accurately, the range of intakes is smaller, and the errors due to spillage and evaporation are relatively greater. Second, many more animals must be tested than for previous research. For example, several hundred mice are required to identify most quantitative trait loci (Bachmanov *et al.*, 1997; Li *et al.*, 2001), and it is likely that several thousand mice will be required to identify taste-related mutations (Schimenti and Bucan, 1998). Third, genetic studies require highly reliable results: Whereas interpretation of most previous research has been based on differences between groups of animals, the results of a single mouse in genetic studies can be

crucial. Inaccurate phenotyping can lead to incorrect localization of quantitative trait loci or a great deal of wasted effort trying to breed, genotype and phenotype the offspring of mice that do not have a genetic anomaly.

We have conducted a series of parametric studies to investigate the extent to which the sensitivity of two-bottle tests is influenced by test duration (Tordoff and Bachmanov, 2002), and the mouse's cage layout (unpublished results), age (Tordoff and Bachmanov, 2001b) and diet (Tordoff *et al.*, 2002). In this paper, we examine whether there are advantages to using more than two bottles to investigate taste preferences.

Several investigators have considered some of the shortcomings and problems with interpretation of two-bottle tests (Warwick and Weingarten, 1996; Sclafani, 2002), particularly their relative advantages and disadvantages compared with one-bottle tests (Dragoin *et al.*, 1971; Grote and Brown, 1971; Elkins, 1973; Klein *et al.*, 1975; Batsell and Best, 1993). There are a few studies that have involved three-bottle tests with a choice among water and two different taste solutions (Zarrow *et al.*, 1960), and one report that a three-bottle test in which one bottle was empty and positioned differently each day had advantages over twobottle tests (Myers and Holman, 1966). Other investigators have conducted experiments involving a choice among six or more bottles (Richter, 1942–1943; Owings *et al.*, 1967; Owings and Lockard, 1968; Goodrick, 1972; Torii *et al.*, 1986). However, all these studies used bottles containing different solutions or different concentrations of the same solution. To our knowledge, there has not been a study in which the same taste solution was provided in two or more bottles simultaneously.

In Experiment 1 reported here, we compared the results of a standard two-bottle test (W-S; W = water, S = solution) with those obtained using three-bottle tests, in which the middle bottle contained a taste solution and the other two contained water (W-S-W), or vice versa (S-W-S). There were two reasons why we thought that three-bottle tests might be advantageous. First, some strains of mice have preferences for the left spout when given a two-bottle choice (Tordoff and Bachmanov, 2001a; Bachmanov *et al.*, 2002), so spout position must be switched half way through each test to control for this. By flanking the taste solution with two water bottles, we hoped to provide a measure of intake unbiased by side preferences. Moreover, by comparing intakes from the two outer tubes, we could measure the side preference. Second, preference scores obtained from two-bottle tests are generally confined to a range of 50%. That is, preferences span from total avoidance (0%) to indifference (50%) for disliked taste solutions, or from indifference (50%) to strong preference (100%) for liked ones. With the three-bottle choice, the range is 67% because preferences span from total avoidance (0%) to indifference (67%) for disliked taste solutions in the S-W-S arrangement, or from indifference (33%) to strong preference (100%) for liked ones in the W-S-W arrangement. The availability of a larger range provided by the appropriate combination of three bottles might allow more discrimination among values, and thus a more sensitive test.

We found in Experiment 1 that mice drank significantly more taste solution when given the S-W-S arrangement than the W-S or W-S-W arrangement of drinking tubes. In Experiment 2a, we examined whether this relationship between the number of available drinking tubes and increased intake generalized to larger numbers. We compared two-bottle tests with tests involving six bottles in which one, two, three, four or five contained 75 mM NaCl and the rest contained water. We were also concerned that differences in 'intake' seen in Experiments 1 and 2a could in fact be due to spillage. Consequently, in Experiment 2b, we repeated parts of Experiment 2a so that we could measure both volume intakes and the number of licks, which is a measure of ingestion that is independent of spillage.

Materials and methods

Subjects and maintenance

All three experiments used separate groups of 15 or 16 male mice. Experiments 1 and 2a involved both C57BL/6J (B6) and 129X1/SvJ (129) strains; Experiment 2b involved just the B6 strain. The mice were purchased from The Jackson Laboratory (Bar Harbor, ME) and were ~12 weeks old at the start of testing. They were individually housed in plastic 'tub' cages (26.5 cm × 17 cm ×12 cm) with a stainless steel grid lid, and wood shavings scattered on the floor [see (Tordoff and Bachmanov, 2001b; Bachmanov *et al.*, 2002) for details]. The vivarium was maintained at 23°C on a 12:12 h light/dark cycle with lights off at 7 pm. The mice were fed pelleted Teklad 8604 chow (Harlan, Madison, WI), and had deionized water to drink.

Drinking tubes

During tests, the mice had access to two or more drinking tubes. Throughout this paper we use the terms 'drinking tubes' and 'bottles' synonymously. When referring to the position of the bottles, we use 'left' and 'right' to refer to the bottles from the viewpoint of a mouse standing in front of and facing the spouts. The drinking tubes were fabricated from plastic pipettes, with stainless steel drinking spouts and rubber stoppers. Each spout extended into the cage 25 mm and had a 3.175 mm diameter hole from which the mouse could lick fluids. Specifics of construction of the drinking tubes are available in earlier papers (Bachmanov *et al.*, 1996, 2002) and in detail on the Monell Mouse Taste Phenotyping Project website (Tordoff and Bachmanov, 2001b). The only difference was that in earlier work we usually fastened to the cage lid a 60 mm × 15 mm metal 'guard' sheet with two holes to accept the drinking spouts, spaced 2 cm apart. This protected the drinking tubes from being chewed by the mice. Instead here, each spout passed through a 2 cm diameter steel washer that rested on the cage lid (Experiment 1), or was protected with a Tygon sheath (Experiment 2).

Experiment 1: comparison of tests using two or three bottles on taste solution preferences

The mice received three series of taste tests. One consisted of the standard two-bottle test with the drinking spouts arranged 2 cm apart (W-S condition). One consisted of three-bottle tests, with the taste solution interposed between two drinking tubes of water (each tube 2 cm from its neighbor; W-S-W condition). The third consisted of three bottle tests, with two drinking tubes containing taste solution flanking a drinking tube containing water (S-W-S condition).

In each series, the following taste solutions were presented along with water: 2 mM saccharin, 50 mM citric acid, 300 μM quinine hydrochloride (QHCl), and 75 mM NaCl. All compounds were purchased from Sigma Chemical Corp. (St Louis, MO), and dissolved in deionized water (deionized water was also used in the W tubes). Taste solutions were made freshly at the beginning of each experiment, and stored in 2 l plastic bottles until needed.

Each series began with a test involving water versus water, and continued with the four taste solution tests given consecutively in the order listed. Each test was 2 days long, so that the whole series lasted 10 days. For tests involving two bottles, the taste solution was always initially presented on the mouse's left and the water on the right, and the position of the drinking tubes was reversed after the 1-day measurement. For tests involving three bottles, the tube positions were not switched. Fluid intakes were measured daily in the middle of the light period, to the nearest 0.1 ml.

The order of taste tests was always the same (water, saccharin, citric acid, QHCl and NaCl) but each series was counterbalanced so that approximately equal numbers of mice received each of the three types of test at the same time.

Each mouse's bedding was changed at the beginning of each series and after the test with 50 mM citric acid. Its body weight was measured at the beginning and end of the experiment.

Experiment 2: six-bottle tests with NaCl as the taste solution

Two experiments were conducted in which mice received various combinations of water and 75 mM NaCl solution in six bottles. In Experiment 2a, 129 and B6 mice received six 48 h tests, with 1–3 days with just water to drink between each test. One test was a standard two-bottle choice (W-S). In the other five, six bottles were always given, with one, two, three, four, or five containing 75 mM NaCl and the others containing water. During the six-bottle tests, the drinking tubes were arranged across the front of the cage (from the mouse's viewpoint) as follows: W-W-S-W-W-W, W-S-W-W-S-W, S-W-S-W-S-W, S-W-S-S-W-S, and S-S-W-SS-S. The order of the tests was counterbalanced so that two or three mice from each strain were tested in each condition at the same time.

In Experiment 2b, B6 mice received four combinations of water and 75 mM NaCl solution, and both volume intakes and every lick were measured. In order to record licks with contact lickometers, each mouse's cage was modified as follows: A 25×8 cm piece of 1/16″ steel sheet was bent into a distorted Z shape and bolted to the cage lid so that it provided a metal platform suspended ~10 cm below the drinking spouts. The mouse had to stand on this steel sheet in order to drink. Each drinking spout was sheathed in Tygon tubing, except for the tip, so that it could not contact the metal cage lid or steel sheet. During tests, a small current (<1 μA) was passed through each spout, such that when the mouse touched the drinking spout it closed a high impedance circuit running to ground. This was detected by a lickometer amplifier (ENV-250C, Med Associates, Georgia, VT) and the shaped output fed into a latched relay (LI-157, Alpha Products, Fairfield, CT), which in turn was connected to a PC computer Bus adapter (Alpha Products, MB-120). Software written in QBASIC scanned the computer input ports once every 100 ms to detect licks.

Because of equipment limitations only two mice were tested at a time. They were given at least 3 days to habituate to the steel platform in their cages. Then, each received in counterbalanced order a two-bottle test and three six-bottle tests. For the six-bottle tests, one, three or five bottles contained 75 mM NaCl and the others contained water. The order of the bottles across the cage was the same as for Experiment 2a. Each test lasted 1 day, and was separated from the following test by 2 or 3 days with only a single bottle of water to drink.

Statistical analyses

All hypothesis testing used a criterion of $P < 0.05$ for statistical significance, but exact probabilities are given below, so readers can use other criteria. Unless otherwise noted, values given in the text and tables are means ±SEMs.

Body weight—In Experiment 1, there was a significant difference between the strains in body weight [B6 = 28 ± 0.3 g, 129 = 30 ± 0.5 g; $t(30) = 3.69$, $P < 0.05$]. However, the difference was small, so we did not consider the effect of body weight on solution intake [see (Bachmanov *et al.*, 1998, 2002) for discussion]. This considerably simplified subsequent analyses. There were no body weight differences between the strains in Experiment 2a (B6 = 30 ± 0.7 g, 129 = 30 ± 0.5 g). The B6 mice in Experiment 2b weighed 28.0 ± 0.5 g.

Experiment 1—For conditions in which only water was presented (i.e. W-W and W-W-W tests), we determined the volume of water ingested by each mouse from the left, middle and right tubes, and its side preference (intake from left tube/total fluid intake). Average daily intakes were analyzed by analyses of variance (ANOVAs). The existence of a significant side preference for a given strain and condition was inferred by comparing the mean preference

with indifference (i.e. 50% in two-bottle tests, 33% in three-bottle tests) using one-sample t-tests. Differences in side preference scores between the strains were analyzed by ANOVA.

For tests involving taste solutions, fluid intakes from each drinking tube for each mouse were collated according to the taste solution and test condition. Solution preference ratios were calculated based on the formula: preference (%) = taste solution intake/(taste solution intake + water intake) × 100. Subsequent analyses were conducted in parallel, using as dependent variables taste solution intakes, water intakes, total fluid intakes, and taste solution preferences.

Strain differences were inferred from planned comparisons confined to the type of test (i.e.W-S, W-S-W, or S-W-S). Differences between types of tests were inferred from planned comparisons based on total intakes of each fluid (i.e. with intakes from the two tubes containing the same fluid summed in three-bottle tests).

Taste test sensitivity—We have developed a method to estimate the sensitivity of a taste test, based on the assumption that this is reflected by the difference between B6 and 129 strains. The method takes into account both the magnitude of the difference between strain means and the within-strain variability. A detailed justification is provided elsewhere (Tordoff and Bachmanov, 2002; Tordoff *et al.*, 2002). In effect, the approach is to consider the 129 mice as 'outliers' of the B6 group, calculate for each test condition how many standard deviations each 129 mouse differs from the B6 mean, and then compare the values obtained from each test condition. To do this, we normalized all data from each test to the B6 group mean. That is, for each test, we calculated *z* scores, based on the mean of the B6 group and the average standard deviations of the B6 and 129 groups. The normalized values of the 129 mice were collated for each of the three test types then compared using one-way ANOVAs for each of the four taste solutions [see (Tordoff and Bachmanov, 2002) for additional justification].

Experiment 2—The results of Experiment 2a were analyzed by two-way ANOVAs with factors of strain and test type. Separate analyses were conducted for water intakes, NaCl intakes, total fluid intakes, and NaCl preference scores. Post hoc comparisons between individual pairs of means were conducted using LSD tests. The results of Experiment 2b were analyzed by one-way ANOVAs (there was no strain factor because only one strain was tested) using the same dependent variables but also water spout licks, NaCl spout licks, and NaCl preference scores derived from licks (licks of spouts containing NaCl/licks of all spouts).

Results

Experiment 1: comparison of tests using two or three bottles on taste solution preferences

Side preferences during tests with two or three bottles of water—During the test with two bottles of water, the 129 mice drank significantly more from the left than right spout whereas the B6 mice drank similar volumes from each spout (Table 1; strain × side interaction [$F(1,30) = 4.13, P = 0.05$; 129 preference, one sample *t*-test comparison with 50%, $t(15) = 2.99, P = 0.009$]). There was no difference in total water intake between the two strains.

Each mouse was tested twice with three bottles of water (once at the beginning of the W-S-W series, and once at the beginning of the S-W-S series). The results of these two tests were almost identical and so we present and analyze average intakes of these two tests here. The 129 mice drank significantly more water from the left spout and significantly less from the middle spout than did the B6 strain, $F(2,60) = 11.2, P < 0.0001$. The two strains had similar intakes from the right spout. The proportion of total intake consumed by the 129 strain from each spout differed significantly from 33% [the expected value; left, $t(15) = 5.12, P = 0.0001$; middle, $t(15) = 6.47, P < 0.0001$, right, $t(15) = 2.61, P = 0.020$]. However, the B6 strain drank from each tube

with preferences that were statistically indistinguishable from 33%. Intakes from the right tube and total intakes (from all three tubes) were similar for the two strains (Table 1).

Tests with taste solutions—Over the three types of test combined, B6 mice had significantly higher preferences for all four taste solutions tested than did 129 mice (Table 2). The only tests in which there was an interaction between strain and test type involved saccharin, and this was apparently due to differences in the volume of water ingested by the two strains (Figure 1, Table 2): for 129 mice, intakes of water differed significantly between each of the three test types but for B6 mice there were no differences.

For each of the four taste solutions, solution intakes and preferences were significantly higher in the S-W-S test than the W-S-W test, and the reciprocal difference was present for water intakes (Figure 1; Table 2). Intakes and preferences in two-bottle tests were intermediate between those of the W-S-W and S-W-S tests, and the differences between two and three-bottle tests were often significant (W-S > W-S-W for intake and preference of saccharin, citric acid, NaCl and QHCl; W-S < W-S-W for water intake with saccharin available; W-S < S-W-S for water intake with NaCl or saccharin available; W-S < S-W-S for intake of NaCl and saccharin and saccharin preference).

Whether mice had two or three bottles had no influence on total fluid intakes when the test solution was citric acid or NaCl. However, for tests involving saccharin or QHCl, the mice drank significantly more total fluid in the S-W-S than W-S condition. With saccharin available, they drank significantly more total fluid in the S-W-S condition than W-S-W condition. Other comparisons were not significant.

Comparison of intakes from the two outer tubes in three-bottle tests provided evidence of side preferences. The 129 mice had significant preferences for the left spout in all W-S-W tests (range $64 \pm 7\%$ – $68 \pm 7\%$), and the S-W-S tests with saccharin ($74 \pm 6\%$) or NaCl ($67 - 6\%$), but not the S-W-S tests with citric acid or QHCl (perhaps because the intakes of these fluids were so low). Side preferences of the B6 mice did not differ significantly from 50% except during the S-W-S test with NaCl ($62 \pm 5\%$).

Comparison of test sensitivity—There were significant differences in the sensitivity of the tests to discriminate 129 from B6 mice for all taste solutions except citric acid [saccharin, $F_{(2,30)} = 10.6, P = 0.0003$; citric acid, $F_{(2,30)} = 1.43, P = 0.23$; QHCl, $F_{(2,30)} = 3.35, P = 0.049$; NaCl, $F_{(2,30)} = 3.67, P < 0.038$]. Differences between individual tests are shown in Table 3.

Experiment 2a: six-bottle tests with NaCl as the taste solution

Intake of 75 mM NaCl and water by both strains depended on how many drinking tubes of each fluid they had available (Figure 2). There was a clear and constant strain difference, with the B6 mice drinking more NaCl than the 129 mice in every test [Figure 2; $F_{(1,28)} = 17.0, P = 0.0003$]. There were no significant differences between the strains in water intake, and no interactions between strain and test type for any measure. That is, the differences between 129 and B6 mice were similar for each test condition. Because of this, the remainder of this section focuses on the results from both strains combined.

NaCl intake (from all NaCl-containing drinking tubes) was significantly higher in six-bottle tests with four or five NaCl bottles than in those with one or two NaCl bottles. NaCl intakes during tests with three NaCl bottles were intermediate between, and did not differ significantly from the two extremes (Figure 2). NaCl intakes in the two-bottle test were similar to those of the six-bottle tests involving two or three NaCl bottles; they were significantly greater than those in the six-bottle test with one NaCl bottle and significantly less than those in the six-bottle tests with four or five NaCl bottles, $F_{(5,140)} = 30.8, P < 0.00001$.

Water intakes showed a reciprocal pattern to NaCl intakes, $F(5,140) = 34.5$, $P < 0.00001$, with the same general distinctions among the six-bottle tests (one NaCl bottle > two NaCl bottles > three or four NaCl bottles > five NaCl bottles). Water intakes in the two-bottle test were similar to those in the six-bottle tests with three or four NaCl bottles. Total fluid intakes also differed significantly among the six test types, $F(5,140) = 9.65$, $P < 0.00001$. Total intakes during the two-bottle test were significantly less than during all five of the six-bottle tests. In addition, total intakes during the tests with one or four NaCl bottles were significantly less than those with two, three or five NaCl bottles.

NaCl preference scores for the two strains combined were $71 \pm 4\%$ for the two-bottle test, and for the one-, two-, three-, four- and five-NaCl six-bottle tests, respectively, $50 \pm 2\%$, $65 \pm 2\%$, $78 \pm 2\%$, $81 \pm 3\%$ and $94 \pm 1\%$, $F(5,140) = 37.4$, $P < 0.000001$. NaCl preference scores for the two-bottle test did not differ significantly from those observed during the six-bottle tests involving two or three bottles of NaCl. There were significant differences in NaCl preference scores between most of the six-bottle tests (one < two < three or four < five < six bottles of NaCl).

We conducted an informal analysis of the distribution of intakes during the five six-bottle tests. B6 mice drank from each of the six bottles about equally (from left to right; ml/day; 1.0 ± 0.1, 1.4 ± 0.3, 1.2 ± 0.1, 1.4 ± 0.1, 1.1 ± 0.1, 1.1 ± 0.1) whereas 129 mice favored the left bottles (from left to right; ml/day; 1.7 ± 0.1, 1.2 ± 0.1, 1.0 ± 0.1, 0.6 ± 0.1, 0.7 ± 0.1, 0.4 ± 0.1). This does not replicate a report indicating that A/J, C57BL and C3H mice prefer to drink from the bottles at both ends in tests with five or seven water bottles (Goodrick, 1972). However, due to the asymmetric distribution of water and NaCl bottles across the cage, some bottle positions in our experiment delivered NaCl more frequently than they delivered water (see Methods). Thus, comparing between experiments requires caution.

Experiment 2b: six-bottle tests with NaCl as the taste solution—lick measurements

B6 mice tested with 75 mM NaCl drank the salty solution in relation to the number of drinking tubes of it they had available. This was true whether the unit if ingestion was milliliters or licks [Figure 3; volume, $F(3,45) = 26.9$, $P < 0.0001$; licks, $F(3,45) = 11.0$, $P < 0.0001$]. Water intakes and licks showed a reciprocal pattern to solution intakes and licks [volume, $F(3,45) = 32.7$, $P < 0.0001$; licks, $F(3,45) = 24.1$, $P < 0.0001$]. The total volume of fluid ingested did not differ significantly among the four tests, although there was a trend for intakes to be higher when six bottles rather than two bottles were available (total fluid intakes; W-W-W-SW-W = 7.0 ± 0.3 ml, W-S-W-S-W-S = 7.6 ± 0.4 ml, S-S-SW-S-S = 6.8 ± 0.5 ml, W-S = 6.1 ± 0.3 ml). Similarly, there were no differences in the total number of licks made during six- and two-bottle tests (W-W-W-S-W-W = 2544 ± 184, W-S-W-S-W-S = 2627 ± 224, S-S-S-W-S-S = 2435 ± 244, W-S = 2045 ± 291).

The number of licks closely matched the volume ingested (Figure 3). Preference scores for NaCl consumption were almost identical whether volume or licks was the measure of ingestion (volume for the one, three and five-bottle conditions, $37 \pm 14\%$, $57 \pm 4\%$ and $79 \pm 4\%$, respectively; licks for the one, three and five-bottle conditions, $31 \pm 6\%$, $58 \pm 4\%$ and $82 \pm 4\%$, respectively).

Discussion

Influence of availability on intake

The results suggest that the number of bottles of taste solution and water has a marked influence on taste solution preference scores. In Experiment 1, the standard two-bottle test (W-S) was compared with two forms of three-bottle test (W-S-W and S-W-S). For each of the four taste

solutions examined, taste solution preferences were highest in the S-W-S test, intermediate in the W-S test, and lowest in the W-S-W test. In Experiments 2a and 2b, the standard two-bottle test was compared with various forms of six-bottle test, using 75 mM NaCl as the taste solution. In both experiments, there was a strong positive linear relationship between the number of bottles of NaCl available and NaCl consumption (both intake and preference). Preference scores during tests with three bottles of NaCl and three of water were similar to those seen with the two-bottle test. We conclude from these findings that the relative availability of taste solution and water determines taste solution preference scores.

These findings led us to conduct several other studies, some of which have already been published. We have shown that: (i) 129 mice, which are normally considered an alcoholavoiding strain, are indifferent to alcohol when given five bottles of 10% alcohol and one of water (Tordoff and Bachmanov, 2003). (ii) B6 mice, which are used as a model of high alcohol preference, avoid it (i.e. show significantly less than a 50% preference for alcohol) when given one bottle of 10% alcohol and five bottles of water (Tordoff and Bachmanov, 2003). (iii) Rats given five bottles of 32% sucrose and one of water become heavier and fatter faster than rats given one bottle of 32% sucrose and five of water (Tordoff, 2002). (iv) Nearly all rats thrive if given cups of each of the three macronutrients but many fail to thrive (because of inadequate protein intakes) if given the three macronutrients plus three extra cups of carbohydrate or fat (Tordoff, 2002). (v) Rats that thrive when given extra cups of fat or carbohydrate eat more of these nutrients than do rats given one cup of each macronutrient only (Tordoff, 2002). The implications and potential mechanisms underlying these findings have been discussed elsewhere (Tordoff, 2002) and so we will not repeat this here, except to mention their main implication: availability has a major influence on choice, even to the extent that it can override genetic or physiological controls. The present results with exemplar sweet, sour, bitter and salty solutions reinforce our conclusions that the 'availability effect' is a general phenomenon that probably applies to all foods and drinks.

The results found here reinforce the obvious truth that preferences observed in the laboratory may have little or no bearing on preferences observed in real life. It is rare that an animal encounters only two food choices in the wild. There are ecological advantages for omnivores to ingest from several sources. Being familiar with many food sources may increase subsequent survival if a single food source becomes unavailable. Consuming small amounts of several sources may also prevent depletion of a source by allowing nibbled on plants to regrow and animal prey populations to regenerate. Moreover, by distributing intake over several sources, the animal can reduce the chances of fatal poisoning. Consumption of a poisonous source can be sufficiently low that little toxic exposure occurs before taste aversion conditioning produces avoidance. In the laboratory, it is important to simplify and control so that interpretable conclusions can be made but we suspect that in the case of food choice experiments, oversimplification has led to the failure to observe a major component of choice, the effect of availability.

The problem of spillage in tests with multiple bottles

A serious concern about interpreting the results of tests involving more than one bottle is that spillage (including evaporation) artificially elevates 'intakes' because each bottle is a separate, additive source of spillage. This is particularly true for experiments where intakes are low, such as those involving mice given strongly disliked test solutions. In these cases the contribution of spillage to 'intake' is relatively greater. Several lines of evidence suggest that spillage contributes to the results found here, but cannot account for them.

First, all things being equal, we would expect spillage to be related to the number of bottles present. Consistent with this, there were significant differences between total fluid 'intakes' in the two- versus three-bottle tests for two of the five solutions (including the water only test)

presented in Experiment 1, and in two- versus six-bottle tests in Experiment 2a but not 2b. If we assume that 'actual' total fluid intakes are unaffected by availability then any differences in measured total fluid intakes can be attributed to spillage. This allows estimation of the average spillage per drinking tube. In the worst case, Experiment 2a, total fluid intakes during the two-bottle tests were 5.5 ± 0.3 ml/day and during the six-bottle tests (combined) were 6.4 ± 0.2 ml/day, which is ~0.2 ml/day additional 'intake' (i.e. spillage) for each additional drinking tube. In contrast, the difference in the six-bottle conditions in NaCl intake when one bottle of NaCl was available (3.1 ± 0.2 ml/day) or five bottles available (5.8 ± 0.2 ml/day) was ~0.7 ml/day per drinking tube, which is over three times larger than the spillage estimate. Of course, this calculation most likely overestimates the contribution of spillage because (i) it is based on the largest difference in total intakes recorded, and (ii) it assumes that 'actual' fluid intakes are unaffected by the number of available bottles. However, providing several bottles of water and/or NaCl may increase actual intakes, as it does with 32% sucrose (Tordoff, 2002).

Second, in several previous studies we have measured the change in fluid volume or weight of drinking tubes placed on empty cages. In work using methods virtually identical to those used in Experiment 2a, we measured fluid spillage in 24 cages with two or six drinking tubes at 2-day intervals over a 24-day period (Tordoff and Bachmanov, 2003). There was no reliable difference between spillage from tubes in cages containing two or six drinking tubes. We found that spillage for the 1476 valid measurements was 0.145 ± 0.275 (SD) ml/day per drinking tube. This is probably an underestimate because spillage is likely to be higher in cages containing a mouse, but it is clearly in line with the ~0.2 ml/day per drinking tube value calculated from the worst case difference in total intakes observed in Experiment 2a.

Third, in Experiment 2b we used the number of licks as well as the volume ingested as measures of intake, and found remarkably similar patterns of response, even though spillage does not affect lick rates. In previous research, other 'spillage independent' measures of ingestion (or its consequences) have also revealed the effect of availability. This includes the effect of five- versus one-bottle of sucrose to increase body weight and fat carcass content (Tordoff, 2002), five- versus one-bottle of 10% alcohol to increase blood alcohol levels (Tordoff and Bachmanov, 2003), and changes in intake related to the availability of extra cups of solid foods [where spillage can be accurately measured (Tordoff, 2002)]. Taken together, we suspect that spillage accounts for 20–30% of the magnitude of the effects seen with taste solutions in the present experiments with mice, and less in previous experiments using rats.

Two-, three- or six-bottle tests: which are best?

Earlier studies have compared the sensitivity of one- versus two-bottle tests, and concluded that for most uses, the two-bottle test is most sensitive (Dragoin *et al.*, 1971; Grote and Brown, 1971; Elkins, 1973; Klein *et al.*, 1975)—but see (Batsell and Best, 1993). The present results raise the question of whether three or more bottles are better. We have argued elsewhere that the sensitivity of taste tests can be estimated from the magnitude of the difference between B6 and 129 strains (Tordoff and Bachmanov, 2002; Tordoff *et al.*, 2002). We found here that there were significant differences in test sensitivity between two- and three-bottle tests in every case except for citric acid (for which there were nonsignificant trends in the expected direction). The most sensitive tests were the S-W-S tests with disliked substances (QHCl and citric acid) and W-S-W tests with liked ones (saccharin and NaCl). Generally, the least sensitive test was the opposite form of three-bottle test. These findings most likely reflect the greater range of preferences available in the appropriate three-bottle tests (i.e. 0–67% for disliked substances in the S-W-S test and 33–100% for liked substances in the W-S-W test; see Introduction).

Another reason to favor three- over two-bottle tests involves the issue of side preferences. Controlling for side preferences in two-bottle tests by switching bottle positions half-way through the test increases spillage and response variability. With three-bottle tests there is no

need to switch the positions of the outer bottles (it would be pointless because they both contain the same fluid). Moreover, comparison of intakes from the two outer bottles provides a measure of side preferences that compares well with those found in two-bottle tests. The increased sensitivity and no need to switch tube positions may make it feasible to conduct three-bottle tests in 24 h, rather than the minimum of 48 h required for two-bottle tests. This can lead to considerable savings in time, effort and per diem housing costs.

There are also several disadvantages of three-bottle tests. They require more equipment. The larger number of drinking tubes raises issues with increased spillage (see above), although this is offset somewhat because the drinking tubes are not disturbed in order to switch positions during the test, as they are for two-bottle tests. In order to take advantage of the increased sensitivity, it is necessary to know whether the compound to be tested will increase or decrease solution preference so that the appropriate form of three-bottle test can be used (W-S-W or S-W-S, respectively). Finally, there is a great deal of inertia involved with changing from a method used for over 70 years to something new.

Despite these disadvantages, we believe that the advantages conferred by greater sensitivity and the potential use of shorter tests makes the three-bottle method superior to its two-bottle cousin in many situations. It is tempting to consider whether, by extension, tests involving more than three bottles have even more advantages. However, we do not see any that would override the problems related to the additional equipment, organization, and spillage involved, unless the goal is to stimulate intake of disliked fluids. We conclude that for most purposes, the optimal method to assess intakes and preferences is the three-bottle test.

Acknowledgements

References

Bachmanov AA, Reed DR, Beauchamp GK, Tordoff MG. Body weight, food intake, water intake, and drinking spout side preference of 28 mouse strains. Behav. Genet 2002;32:435–443. [PubMed: 12467341]

Bachmanov AA, Reed DR, Ninomiya Y, Inoue M, Tordoff MG, Price RA, Beauchamp GK. Sucrose consumption in mice: major influence of two genetic loci affecting peripheral sensory responses. Mamm. Genome 1997;8:545–548. [PubMed: 9250857]

Bachmanov AA, Tordoff MG, Beauchamp GK. Ethanol consumption and taste preferences in C57BL/6ByJ and 129/J mice. Alcohol Clin. Exp. Res 1996;20:201–206. [PubMed: 8730208]

Bachmanov AA, Tordoff MG, Beauchamp GK. Voluntary sodium chloride consumption by mice: differences among five inbred strains. Behav. Genet 1998;28:117–124. [PubMed: 9583237]

Batsell WR, Best MR. One bottle too many? Method of testing determines the detection of overshadowing and retention of taste aversions. Anim. Learn. Behav 1993;21:154–158.

Dragoin W, McCleary GE, McCleary P. A comparison of two methods of measuring conditioned taste aversion. Behav. Res. Meth. Instrum 1971;3:309–310.

Elkins RL. Individual differences in bait shyness: effects of drug dose and measurement technique. Psychol. Record 1973;23:349–358.

Goodrick CL. End bottle preferences of inbred mice during alcohol preference and fluid intake multiple-bottle test procedures. Psychon. Sci 1972;28:185–187.

Grote FW, Brown RT. Conditioned taste aversions: two-stimulus tests are more sensitive than one-stimulus tests. Behav. Res. Methods Instrum 1971;3:311–312.

Klein SB, Domato GC, Hallstead C, Stephens I, Mikulka PJ. Acquisition of a conditioned aversion as a function of age and measurement technique. Physiol. Psychol 1975;3:379–384.

Li X, Inoue M, Reed DR, Huque T, Puchalski RB, Tordoff MG, Ninomiya Y, Beauchamp GK, Bachmanov AA. High-resolution genetic mapping of the saccharin preference locus (Sac) and the putative sweet taste receptor (T1R1) gene (Gpr70) to mouse distal Chromosome 4. Mamm. Genome 2001;12:13–16. [PubMed: 11178737]

Myers RD, Holman RB. A procedure for eliminating position habit in preference–aversion tests for ethanol and other fluids. Psychon. Sci 1966;6:235–236.

Owings DH, Haerer HA, Lockard RB. Sucrose intake functions of rat and cockroach for single and six solution presentations. Psychon. Sci 1967;73:125–127.

Owings DH, Lockard RB. Effects of deprivation and presence of food on intake of sucrose solution in six-bottle tests. Psychon. Sci 1968;12:299–300.

Richter CP. Total self-regulatory functions in animals and human beings. Harvey Lectures 1942–1943:63–103.38

Schimenti J, Bucan M. Functional genomics in the mouse: phenotype-based mutagenesis screens. Genome Res 1998;8:698–710. [PubMed: 9685317]

Sclafani A. Flavor preferences conditioned by sucrose depend upon training and testing methods. Two-bottle tests revisited. Physiol. Behav 2002;76:633–644. [PubMed: 12127003]

Tordoff MG. Obesity by choice: the powerful effect of nutrient availability on nutrient intake. Am. J. Physiol. Regul. Integr. Comp. Physiol 2002;282:R1536–R1539. [PubMed: 11959698]

Tordoff, MG.; Bachmanov, AA. Food intakes, water intakes, and spout side preferences. 2001a. http://aretha.jax.org/pub-cgi/phenome/mpdcgi?rtn=studies/details&id=63

Tordoff, MG.; Bachmanov, AA. Monell mouse taste phenotyping project. 2001b. www.monell.org/MMTPP

Tordoff MG, Bachmanov AA. Influence of test duration on the sensitivity of the two-bottle choice test. Chem. Senses 2002;27:759–768. [PubMed: 12438201]

Tordoff MG, Bachmanov AA. Influence of the number of alcohol and water bottles on murine alcohol intake. Alcohol Clin. Exp. Res 2003;27:600–606. [PubMed: 12711922]

Tordoff MG, Pilchak DM, Williams JA, McDaniel AH, Bachmanov AA. The maintenance diets of C57BL/6J and 129X1/SvJ mice influence their taste solution preferences: implications for large-scale phenotyping projects. J. Nutr 2002;132:2288–2297. [PubMed: 12163677]

Torii, K.; Mimura, T.; Yugari, Y. Effects of dietary protein on the taste preference for amino acids in rats. In: Kare, MR.; Brand, JG., editors. Interaction of the chemical senses with nutrition. Academic Press; Orlando, Florida: 1986. p. 45-69.

Warwick ZS, Weingarten HP. Flavor-postingestive consequence associations incorporate the behaviorally opposing effects of positive reinforcement and anticipated satiety: implications for interpreting two-bottle tests. Physiol Behav 1996;60:711–715. [PubMed: 8873241]

Zarrow MX, Aduss H, Denison ME. Failure of the endocrine system to influence alcohol choice in rats. Q. J. Stud. Alcohol 1960;21:400–413. [PubMed: 13787952]
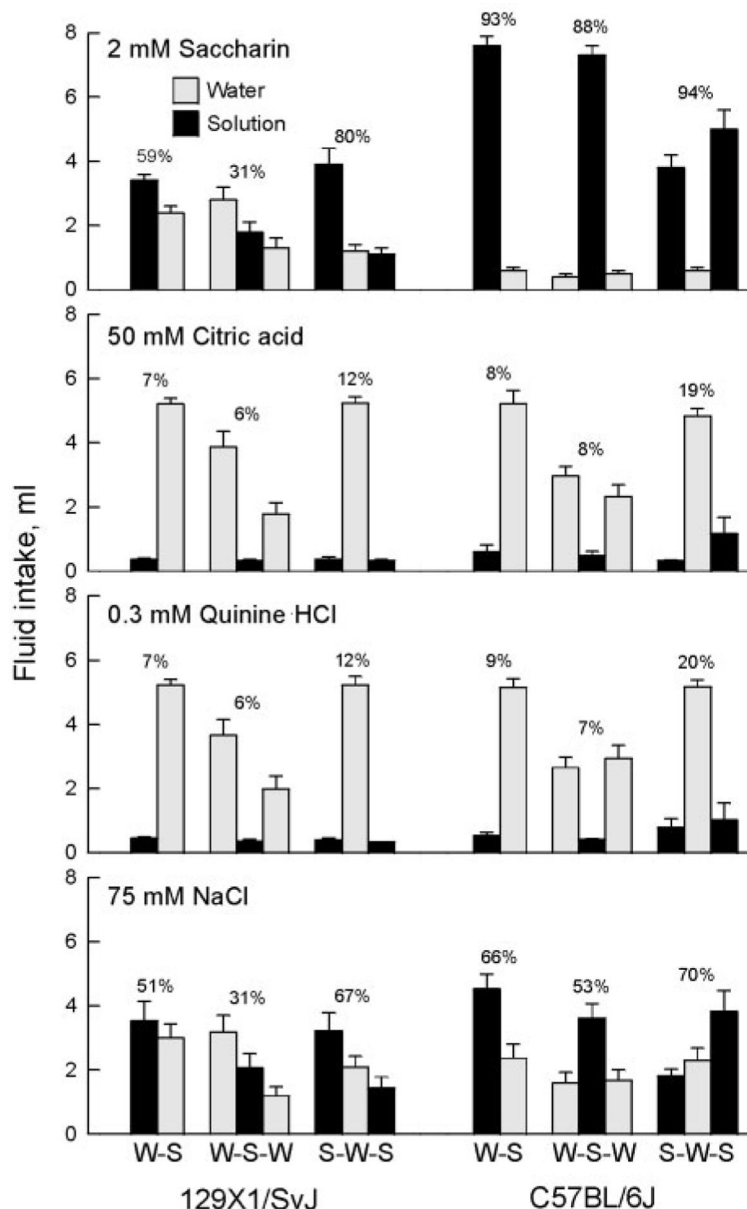
Figure 1.
Intakes of water and taste solutions by 129X1/SvJ and C57BL/6J mice tested with one bottle of water and one of taste solution (W-S), two bottles of water and one of taste solution (W-S-W), or one bottle of water and two of taste solution (S-W-S). Values above columns are taste solution preferences in relation to total fluid intake. For the three-bottle tests, the results from individual drinking tubes are arranged left–middle–right. Note that the 129 mice generally drink more from the left than right drinking tube. Note also that total solution intakes and solution preferences (from both drinking tubes combined) were significantly higher in the S-W-S than W-S-W conditions.

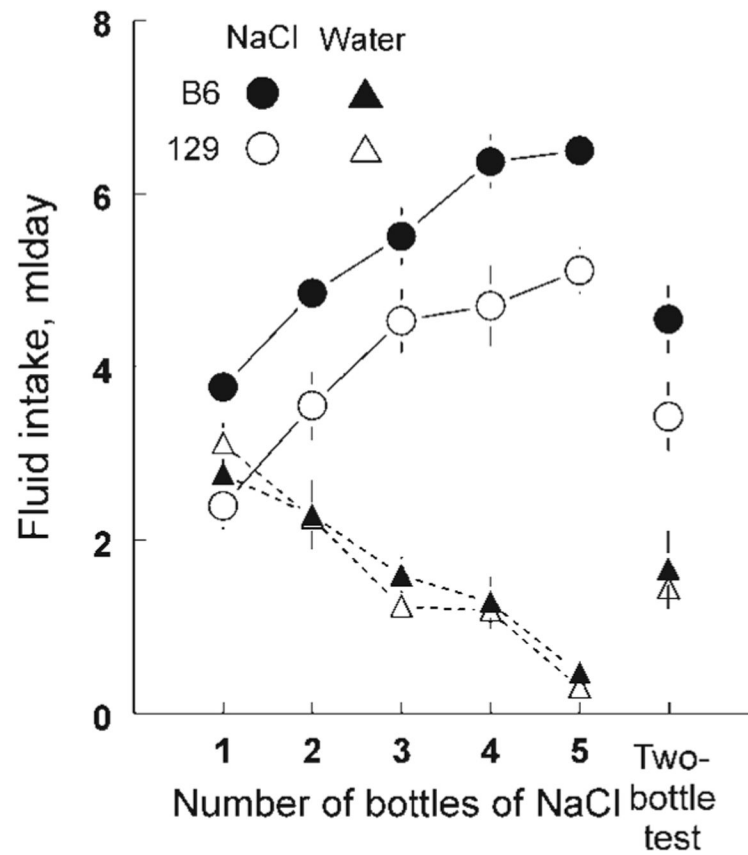**Figure 2.**
Daily intake of 75 mM NaCl and water by groups of 15 C57BL/6J (B6) and 129X1/SvJ (129) mice given a two-bottle test (right) or six-bottle tests with access to one, two, three, four or five bottles of NaCl and the rest containing water.
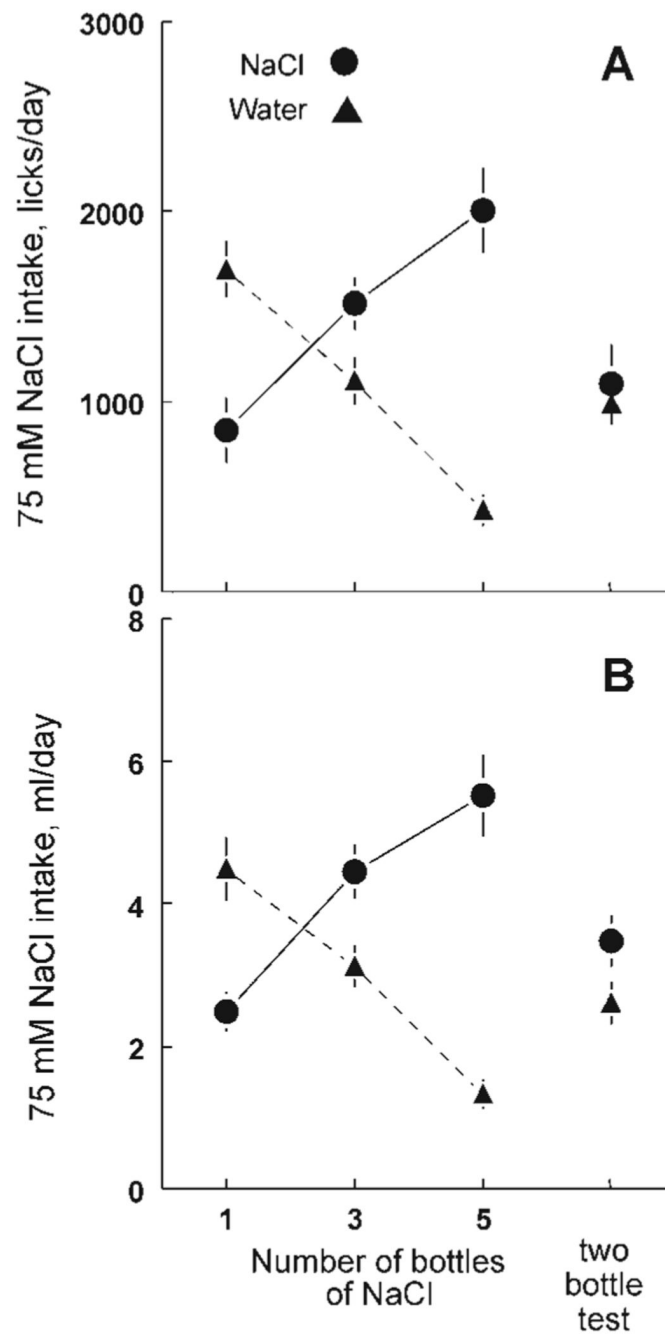
Figure 3.
Mean ± SE number of spout licks **(A)** and volume intakes **(B)** of 75 mM NaCl and water by
C57BL/6J mice given a regular two-bottle test (right of each panel) or six-bottle tests with
access to one, three or five bottles of NaCl and the rest containing water.

Table 1

Intake and spout side preference of B6 and 129 mice during tests with two or three bottles of water in Experiment 1

Test type and source of water	Intake (ml)		Preference (%)	
	129	**B6**	**129**	**B6**
Two-bottle test				
Left	3.5 ± 0.3	3.2 ± 0.3	$64 \pm 5^{*\dagger}$	52 ± 4
Right	$2.0 \pm 0.3^{*}$	2.9 ± 0.3	$36 \pm 5^{*\dagger}$	48 ± 4
Total	5.5 ± 0.2	6.0 ± 0.5	$-$	$-$
Three-bottle test				
Left	$3.5 \pm 0.4^{*}$	2.1 ± 0.2	$58 \pm 5^{*\dagger}$	35 ± 4
Middle	$1.0 \pm 0.1^{*}$	2.2 ± 0.3	$18 \pm 2^{*\dagger}$	36 ± 4
Right	1.3 ± 0.2	1.8 ± 0.2	$24 \pm 4^{\dagger}$	30 ± 2
Total	5.8 ± 0.2	6.1 ± 0.3	$-$	$-$

*
$P < 0.05$ relative to B6 strain

†
$P < 0.05$ relative to indifference (i.e. 50% preference for two-bottle test or 33% preference for three-bottle test). Values for three-bottle test are average of two tests for each mouse.

Table 2

Results of ANOVAs comparing B6 and 129 mice given three types of test with two or three drinking spouts (Experiment 1)

Source of variance and dependent variable	2 mM saccharin	50 mM citric acid	0.3 mM QHCl	75 mM NaCl
Effect of strain (d.f. = 1 and 30)				
Water intake	147.9****	0.66	0.04	1.00
Solution intake	228.5****	3.30	3.24	3.48
Total fluid intake	54.0****	0.11	0.64	1.44
Solution preference	315.3****	4.74*	4.65*	5.37*
Effect of test type [a] (d.f. = 2 and 60)				
Water intake	39.7****	3.25*	3.58*	16.9****
Solution intake	36.5****	6.62***	6.51***	17.4****
Total fluid intake	6.4***	0.93	3.15*	2.39****
Solution preference	46.6****	18.1****	17.4****	21.9****
Strain × test type interaction (d.f. = 2 and 60)				
Water intake	23.0****	0.94	0.00	2.68
Solution intake	2.2	1.32	2.35	0.33
Total fluid intake	3.2*	0.43	1.73	0.92
Solution preference	28.1****	1.96	2.75	3.02

Values in the body of the table are F scores.

*
$P < 0.05$

**
$P < 0.01$

***
$P < 0.005$

****
$P < 0.00001$.

[a] The three test types were W-S-W = three bottle test with a bottle of taste solution flanked by two bottles of water, W-S = two-bottle test, S-W-S = bottle of water flanked by two bottles of taste solution.

Table 3

Influence of drinking spout layout on the variability of preference scores: deviation of B6 and 129 mice from the mean value of the B6 strain

		2 mM saccharin	50 mM citric acid	0.3 mM QHCl	75 mM NaCl
Three bottle (W-S-W)	B6	0.00 ± 0.07	0.00 ± 0.35	0.00 ± 0.14	0.00 ± 0.25
	129	4.99 ± 0.43^c	0.61 ± 0.15	0.26 ± 0.36^a	0.88 ± 0.25^b
Two-bottle (W-S)	B6	0.00 ± 0.09	0.00 ± 0.33	0.00 ± 0.26	0.00 ± 0.20
	129	3.99 ± 0.41^b	0.45 ± 0.17	0.30 ± 0.24^a	0.61 ± 0.30^b
Three bottle (S-W-S)	B6	0.00 ± 0.09	0.00 ± 0.39	0.00 ± 0.41	0.00 ± 0.23
	129	2.13 ± 0.41^a	0.78 ± 0.11	0.82 ± 0.09^b	0.13 ± 0.27^a

Values in the body of the text show mean $\pm$ SEM of z scores, based on the mean of the B6 group and the average standard deviations of the B6 and 129 groups. Thus, for example, in the 'Two-bottle (W-S)' condition with 2 mM saccharin, the mean of the 129 mice differed from the mean of B6 mice by 3.99 standard deviations. In all cases, the B6 mean was higher than the 129 mean. All means for the B6 strain are 0.00 because z scores are based on this mean. A significant difference in the ANOVA (shown in text) signifies that there were differences among the tests in their ability to distinguish B6 from 129 mice (the ANOVA was not significant for the results with citric acid). Values in the same column with the same superscript did not differ significantly from each other ($P < 0.05$, LSD test).