tetradecanol) are separated and detected, quantitation is hindered by poor solubility and the formation of micelles.

Applications. Since a prerequisite for PAD reactivity is adsorption of the analyte, the presence of other surface-adsorbable substances, as well as electroactive compounds, can act as interferences. Therefore, general selectivity is achieved via chromatographic separation prior to PAD. This conclusion does not preclude additional selectivity from control of detection parameters.

The assay for alcohols was applied to several matrices to illustrate the analytical utility of the procedure. Separation on an ion-exclusion column with direct detection is illustrated in Figure 8 for various aliphatic alcohols and polyalcohols in toothpaste (A), liquid cold formula (B), brandy (C), and wine cooler (D). The selectivity for alcohols in acidic media at a Pt electrode contributes to decreased time for sample preparation and simplified chromatograms.

The versatility of separations on mixed-mode ion-exchange columns with selective detection is illustrated in Figure 9 by the simultaneous detection of ionic and neutral species in a pharmaceutical preparation. This experiment utilizes a UV detector and PAD in series after a PCX-500 column. Under acidic conditions, the cephalosporin antibacterial consists of a cation (i.e., cefazolin) and neutral and anionic compounds (i.e., 1,6-hexanediol, 1,4-cyclohexanediol, and *p*-toluenesulfonic acid). The neutral and anionic compounds are separated by the reversed-phase character of the column, while the cationic compound is separated by a combination of cation exchange and reversed-phase mechanism. Figure 9 shows that the *p*-toluenesulfonic acid and the cefazolin are both detected by UV at 254 nm (A), and the two diols, which do not have a chromophore, are easily detected by PAD (B). In addition, the cefazolin has a PAD signal, which has a higher limit of detection than UV, but may be utilized for added selectivity.

LITERATURE CITED

(1) Kissinger, P. T. In *Laboratory Techniques in Electroanalytical Chemistry*; Kissinger, P. T., Heineman, W. R., Eds.; Marcel Dekker: New York, 1984; pp 631–32.
(2) Adams, R. N. *Electrochemistry at Solid Electrodes*; Marcel Dekker: New York, 1969.
(3) Gilman, S. In *Electroanalytical Chemistry*; Bard, A. J., Ed.; Marcel Dekker: New York, 1967; Vol. 2, pp 111–92.
(4) Fleet, B.; Little, C. J. *J. Chromatogr. Sci.* **1974**, *12*, 747.
(5) Van Rooijen, H. W.; Poppe, H. *Anal. Chim. Acta* **1981**, *130*, 9.
(6) Hughes, S.; Meschi, P. L.; Johnson, D. C. *Anal. Chim. Acta* **1981**, *132*, 1.
(7) Johnson, D. C.; LaCourse, W. R. *Anal. Chem.* **1990**, *62*, 589A.
(8) Nachtmann, F.; Budna, K. W. *J. Chromatogr.* **1977**, *136*, 279.
(9) Jupille, T. *J. Chromatogr. Sci.* **1979**, *17*, 160.
(10) Beden, B.; Cetin, I.; Kahyaoglu, D.; Takky, D.; Lamy, C. *J. Catal.* **1987**, *104*, 37.
(11) Ocon, P.; Alonso, C.; Celdran, R.; Gonzalez-Velasco, J. *J. Electroanal. Chem.* **1986**, *206*, 179.
(12) Slingsby, R. W.; Rey, M. *J. Liq. Chromatogr.* **1990**, *13* (1), 107.
(13) LaCourse, W. R.; Jackson, W. A.; Johnson, D. C. *Anal. Chem.* **1989**, *61*, 2466.

RECEIVED for review July 20, 1990. Accepted October 8, 1990. The financial support of Dionex Corporation is acknowledged with gratitude.

Statistical Treatment for Rejection of Deviant Values: Critical Values of Dixon's "*Q*" Parameter and Related Subrange Ratios at the 95% Confidence Level

David B. Rorabacher

Department of Chemistry, Wayne State University, Detroit, Michigan 48202

Critical values at the 95% confidence level for the two-tailed *Q* test, and related tests based upon subrange ratios, for the statistical rejection of outlying data have been interpolated by applying cubic regression analysis to the values originally published by Dixon. Corrections to errors in Dixon's original tables are also included. The resultant values are judged to be accurate to within ±0.002 and corroborate the fact that corresponding critical values published in recent statistical treatises for analytical chemists are erroneous. It is recommended that the newly generated 95% critical values be adopted by analytical chemists as the general standard for the rejection of outlier values.

Analytical chemists depend upon the generation and interpretation of precise experimental data. As a result, they are especially cognizant of the value of statistics in data treatment, and a number of statistical treatises have recently been published that are specifically written for the professional analytical chemist (*1*). Included in each of these publications is a brief section dealing with tests for the rejection of grossly deviant values (outliers). Although many statistical tests have been proposed to deal with this topic [Barnett and Lewis (*2*) discuss 47 different equations designed for this purpose], it is interesting to note that these treatises, as well as essentially all analytical chemistry textbooks published in the U.S. during the past decade (*3*), have settled on the use of Dixon's *Q* test (and variants thereof) (*4*) as the primary method for testing for the rejection of outlying values.

Each of the recent statistical treatises written for analytical chemists has attempted to include critical values of *Q* for the 95% confidence level, values that were not included in Dixon's publications. However, not only do the 95% confidence values differ in each treatise but all compilations contain significant errors. The most legitimate set of 95% values is that presented by Miller and Miller ($4 \le n \le 10$) (*1a*), which they attribute to King (*5*), but no such values are listed in King's article, and the values of Miller and Miller differ by amounts varying from 0.002 to 0.007 from the 95% values presented in the current manuscript. Anderson (*1b*) describes the *equations* corresponding to the *two-tailed* tests for Dixon's parameters designated as r_{10} (for $3 \le n \le 10$), r_{11} (for $8 \le n \le 10$), and r_{21} (for $11 \le n \le 13$) and purportedly lists critical values for the 90%, 95%, and 99% confidence levels for these sample sizes, but the values actually listed in his table are Dixon's values for *one-tailed* tests. Thus, as applied to two-tailed tests, Anderson's confidence levels should be labeled 80%, 90%, and 98% (vide infra). Caulcutt and Boddy (*1c*), while describing *only the equation for the Q* (i.e., r_{10}) *ratio*, accurately list both

0003-2700/91/0363-0139$02.50/0

95% and 99% confidence level critical values for this ratio for five sample sizes, $3 \leq n \leq 7$; the values given in their Table G (p 248) for $8 \leq n \leq 12$ and for $13 \leq n \leq 40$ are clearly *not* Q values but are somewhat similar to (though not identical with) the critical values for r_{11} and r_{21}, respectively. However, *the equations for these latter parameters are not even mentioned in this treatise.*

Faced with this confusing and conflicting array of critical values for the popular Q test and Dixon's related tests based on various subrange ratios, it seems likely that many erroneous statistical inferences are being made daily in analytical chemistry laboratories throughout the world. To counteract this possibility and to extend the utility of Dixon's parameters for two-tailed tests, the current paper provides a reliable set of critical values at the 95% confidence level for the Q parameter and for all of the related subrange ratios proposed by Dixon. It also corrects several errors found in Dixon's original published tables. At the same time, it is intended to alert analytical chemists to the problems that exist in current treatises and, hopefully, may serve as a useful source for future textbook and treatise authors dealing with this topic.

CRITERIA FOR REJECTION OF OUTLIERS

Determinate Errors and Outliers. In any set of analytical measurements, one or more values may incorporate a determinate (systematic) error. Such errors may involve either a "locator" error (displacement of the population mean) or a "scalar" error (increase in the population variance) or a combination of the two (*4*). Values that are affected by a determinate error belong to a different population of measurements. Thus, if the presence of a determinate error is known to the individual making the measurements, the values so affected should automatically be rejected, regardless of their magnitude. (For a brief discussion on the detection and elimination of determinate errors in analytical measurements, see ref 3b.) However, in many instances, determinate errors are incurred that are not detected by the experimentalist.

If an undetected determinate error affects only a fraction of the values in the total sample and if the error is sizable relative to the experimental precision of the remaining values, the affected values may be significantly larger or smaller than the values that are not affected by such errors; i.e., they will be observed as *outlier* values. Recognizing this, experimentalists tend to suspect the presence of determinate errors whenever outliers are observed in a set of data, and there is a strong motivation to eliminate such outliers since they significantly affect the sample mean and, therefore, the final estimate of the population mean. In the absence of independent knowledge that determinate errors are responsible for the appearance of outliers, however, the decision to delete deviant values must be based upon reasonable statistical evidence. Some statisticians look askance at any attempt to discard deviant values by means of statistical criteria. Deming, a highly respected authority in the industrial application of statistics, has concluded that "a point is never to be excluded on statistical grounds alone" (*6*). However, the majority of those who have addressed themselves to the statistical treatment of scientific measurements would appear to agree with Parratt's statement that "rejection on the basis of a hunch or of general fear is not at all satisfactory, and some sort of objective criterion is better than none" (*7*). In fact, the statistical treatment of outlier values has received increasing attention in recent decades and the various statistical tests that have been proposed are described and compared in texts (*2, 8*) and an extensive review (*9*) dealing specifically with this topic.

Tests for Rejection of Grossly Deviant Values. Statistical approaches used for identifying values that are affected by determinate errors are based on testing the hypothesis that such values can be attributed to random variation alone within some reasonable level of probability. If the probability of obtaining such outliers is determined to be very small, based on the overall population variance, one may reasonably conclude that this hypothesis is incorrect, i.e., these values cannot be attributed to simple random variation and are likely the result of a determinate error. On this basis, suspected values may be rejected in a statistically legitimate fashion, provided that the confidence level chosen is a reasonable one.

Confidence levels commonly used for tests involving the rejection of data are 99%, 95% or 90%, the higher levels representing the more conservative approach. Operating at a high confidence level reduces the likelihood of rejecting a legitimate value containing no determinate error (i.e., it reduces the so-called α risk or "Type I error"), but it also increases the likelihood of retaining a value that contains a determinate error (i.e., it increases the so-called β risk or "Type II error") (*10*). The consensus of most statisticians is that the former error (rejecting a "good" value) is considered more serious and, therefore, the use of high confidence levels in the rejection of data is greatly preferred. For example, when operating at the 99% confidence level ($\alpha = 0.01$) with a population containing no outliers, only 1% of the good values will be rejected. In contrast, when operating at the 90% confidence level ($\alpha = 0.10$), 10% of the good values will be rejected and, for small sample sizes, the effect upon the estimation of the population mean becomes significant relative to the accuracy desired by most analytical procedures. [However, Dixon has argued that the use of rejection tests at the 90% confidence level is warranted in many cases (*11*).]

Since the use of a 99% confidence level ($\alpha = 0.01$) allows for the rejection of only the most extreme deviations—particularly when applied to small sample sizes—the β risk of retaining a value that does not belong to the general sample population (i.e., a "bad" value) is quite large, and most experimentalists are not satisfied when operating at this level in testing for the rejection of data. Operating at the 95% confidence level ($\alpha = 0.05$) provides a reasonable compromise. If values not rejected at the 95% confidence level are viewed with suspicion, additional measurements are probably warranted. It should be noted that no specific test or choice of confidence level is ideal, however. As Natrella has noted, "the only sure way to avoid publishing any 'bad' results is to throw away all results" (*12*). The converse, of course, is also true—the only sure way to avoid discarding any good results is to retain them all.

A number of tests, such as those based on the "Student's t distribution" (*13*), require independent knowledge of the population standard deviation (σ) and the population mean (μ) or the uncontaminated sample standard deviation (s) and the uncontaminated sample mean ($\bar{X}$) for their application (*2*). In the absence of such independent information, these tests are of limited use since a decision must be made as to whether the suspected values should be included or excluded in the calculation of s and $\bar{X}$. Neither decision is satisfactory as the choice made automatically biases the outcome of the test. [The often-cited "Chauvenet's criterion" for the rejection of a deviant measurement (*2, 7*) is also based on the sample standard deviation and, therefore, poses the same problems when applied to small samples; it has been noted that this criterion is set to reject, on average, half an observation of good data per sample, regardless of the sample size (*8,14*).] For analytical measurements, where σ and μ (and uncontaminated values of s and $\bar{X}$) are generally not known independently, tests that do not require the use of these quantities are greatly preferred.

APPLICATION OF DIXON'S RANGE TESTS

The Q Test and Related Subrange Ratios. In a classic

1950 article (*4*), Dixon investigated the performance of several statistical tests in terms of their ability to reject bad values in data sets taken from Gaussian populations. The tests investigated included both those which require independent knowledge of σ or s and those which do not require such information. Of the tests included in the latter group, Dixon concluded that tests based on ratios of the range and various subranges were to be preferred as a result of their excellent performance and ease of calculation. [Dixon also noted that another test which performs well in screening for outliers is a modified F test, in which the ratio of the standard deviations calculated by including and excluding the suspected deviant value is compared to critical values of F; however, this latter test may be "masked" by a second deviant value.] The range tests, all of which are closely related, include the following (where the values are ordered such that $x_1 < x_2 < ... < x_{n-1} < x_n$):

1. For a single outlier x_1

$$r_{10} = \frac{x_2 - x_1}{x_n - x_1} \quad \left(\text{OR} \quad \frac{x_n - x_{n-1}}{x_n - x_1}\right)$$

2. For outlier x_1 avoiding x_n

$$r_{11} = \frac{x_2 - x_1}{x_{n-1} - x_1} \quad \left(\text{OR} \quad \frac{x_n - x_{n-1}}{x_n - x_2}\right)$$

3. For outlier x_1 avoiding x_n, x_{n-1}

$$r_{12} = \frac{x_2 - x_1}{x_{n-2} - x_1} \quad \left(\text{OR} \quad \frac{x_n - x_{n-1}}{x_n - x_3}\right)$$

4. For outlier x_1 avoiding x_2

$$r_{20} = \frac{x_3 - x_1}{x_n - x_1} \quad \left(\text{OR} \quad \frac{x_n - x_{n-2}}{x_n - x_1}\right)$$

5. For outlier x_1 avoiding x_2 and x_n

$$r_{21} = \frac{x_3 - x_1}{x_{n-1} - x_1} \quad \left(\text{OR} \quad \frac{x_n - x_{n-2}}{x_n - x_2}\right)$$

6. For outlier x_1 avoiding x_2 and x_n, x_{n-1}

$$r_{22} = \frac{x_3 - x_1}{x_{n-2} - x_1} \quad \left(\text{OR} \quad \frac{x_n - x_{n-2}}{x_n - x_3}\right)$$

(The parenthetical equations are designed for testing x_n, the highest value rather than the lowest value, x_1.) In Dixon's notation, the first digit in the subscript of each ratio, r_{ij}, refers to the number of possible suspected outliers on the same end of the data as the value being tested, while the second digit indicates the number of possible outliers on the opposite end of the data from the suspected value. Thus, the ratio r_{10} simply compares the difference between a single suspected outlier (x_1 or x_n) and its nearest-neighboring value to the overall range of values in the sample—in other words, it determines the fraction of the total range that is attributable to one suspected outlier. The other ratios are similarly formulated except that they use subranges that are specifically designed to avoid the influence of additional outliers either on the opposite end of the data (r_{11} and r_{12}), on the same end of the data (r_{20}), or both (r_{21} and r_{22}). Clearly, the latter ratios require larger sample sizes to perform satisfactorily. Dixon subsequently generated critical values for all of these ratios (*15*) for sample sizes of $3 \leq n \leq 30$ and recommended (based on a combination of the relative performance of each ratio and its degree of independence from other outlying values) that, as a general rule, the various ratios be applied as follows (*16*): for $3 \leq n \leq 7$, use r_{10}; for $8 \leq n \leq 10$, use r_{11}; for $11 \leq n \leq 13$, use r_{21}; for $n \geq 14$, use r_{22}.

The r_{10} ratio is commonly designated as "Q" and is generally considered to be the most convenient, legitimate, statistical test available for the rejection of deviant values from a small sample conforming to a Gaussian distribution. (It is equally well suited to larger data sets if only one outlier is present.) The fact that small data sets are common in analytical testing procedures, in combination with the simplicity of this test, accounts for the fact that the Q test is included in nearly all modern statistical treatises and textbooks designed for use in analytical chemistry (*1, 3*).

A few authors (*17*), following Deming's viewpoint (vide supra), object to the rejection of data from any small sample based on statistical tests, claiming that the amount of information available is insufficient to establish the distribution pattern; recommended alternatives include dropping the highest and lowest values (for a sample containing five or more values) or reporting the median. However, this would appear to be overly cautious since most series of repetitive analytical measurements follow a Gaussian distribution (*18*) provided that s is small compared to $\bar{X}$ (thus, the Q test is not applicable to analytical measurements when operating close to the detection limits). Dixon has tested the relative merits of the sample mean and median as an estimator of the population mean under various conditions (*11, 15*) and has concluded that, for the most part, the sample mean (after the rejection of outliers) appears to provide a better approximation than does the median.

In his original calculation of the critical values of the various r criteria (*15*), Dixon was able to obtain exact solutions only for the case where $n = 3$ or 4. Critical values for $n = 5$, 7, 10, 15, 20, 25, and 30 were calculated by using numerical methods. All other values were obtained by interpolation and were generally judged to be accurate within ±0.001.

An irritating feature of the Q test and Dixon's related subrange ratio tests, as they currently exist, is the lack of suitable critical values of Q (and r_{11}, r_{21}, etc.) at the 95% confidence level ($\alpha = 0.05$), since this confidence level is frequently utilized for all other statistical tests. The lack of 95% confidence level values arises from the fact that Dixon generated critical values at several standard probability levels ($\alpha = 0.005$, 0.01, 0.02, 0.05, etc.) corresponding to 99.5%, 99%, 98%, 95%, etc., confidence levels in terms of a *one-tailed* test (*15*). As commonly applied by analytical chemists and other experimentalists, however, these tests are used as *two-tailed* tests (i.e., one is generally interested in testing outlier values at both the upper and lower ends). The probability that an individual good value may lie outside a specified interval in either tail (i.e., on either the high or low end) is twice as large as the probability that it lies outside the chosen interval in only one tail. As a result, the α risk doubles for a two-tailed test (*18*) and the confidence level decreases accordingly. Thus, Dixon's 99.5%, 99%, 98%, and 95% confidence levels translate into 99%, 98%, 96%, and 90% levels, respectively, when considering a two-tailed test, a fact not recognized by some authors (*1b, 3e,m*). Since Dixon did not calculate one-tailed critical values at the 97.5% confidence level, there have been no values available at the 95% level when using these tests as two-tailed tests. Therefore, about half of the current analytical chemistry textbooks (*3a–d,h*) listing valid critical values of Q list the values for the 99%, 96%, and 90% confidence levels, despite the fact that the 96% confidence level is somewhat unorthodox, while the remainder (*3f,g,i–l*) provide only 90% confidence level values. As noted earlier, the treatises that have attempted to present critical values for Q at the 95% confidence level (*1*) have invariably listed flawed values.

[In tests designed to detect the presence of a single outlier, King has argued (*5*) that the effect of running a two-tailed test is *approximately* to double the α risk relative to that for

Table I. Critical Values of Dixon's r_{10} (Q) Parameter As Applied to a Two-Tailed Test at Various Confidence Levels, Including the 95% Confidence Level[a]

	confidence level					
N[b]	80% (α = 0.20)	90% (α = 0.10)	95% (α = 0.05)	96% (α = 0.04)	98% (α = 0.02)	99% (α = 0.01)
3	0.886	0.941	**0.970**	0.976	0.988	0.994
4	0.679	0.765	**0.829**	0.846	0.889	0.926
5	0.557	0.642	**0.710**	0.729	0.780	0.821
6	0.482	0.560	**0.625**	0.644	0.698	0.740
7	0.434	0.507	**0.568**	0.586	0.637	0.680
8	0.399	0.468	**0.526**	0.543	0.590	0.634
9	0.370	0.437	**0.493**	0.510	0.555	0.598
10	0.349	0.412	**0.466**	0.483	0.527	0.568
11	0.332	0.392	**0.444**	0.460	0.502	0.542
12	0.318	0.376	**0.426**	0.441	0.482	0.522
13	0.305	0.361	**0.410**	0.425	0.465	0.503
14	0.294	0.349	**0.396**	0.411	0.450	0.488
15	0.285	0.338	**0.384**	0.399	0.438	0.475
16	0.277	0.329	**0.374**	0.388	0.426	0.463
17	0.269	0.320	**0.365**	0.379	0.416	0.452
18	0.263	0.313	**0.356**	0.370	0.407	0.442
19	0.258	0.306	**0.349**	0.363	0.398	0.433
20	0.252	0.300	**0.342**	0.356	0.391	0.425
21	0.247	0.295	**0.337**	0.350	0.384	0.418
22	0.242	0.290	**0.331**	0.344	0.378	0.411
23	0.238	0.285	**0.326**	0.338	0.372	0.404
24	0.234	0.281	**0.321**	0.333	0.367	0.399
25	0.230	0.277	**0.317**	0.329	0.362	0.393
29	0.227	0.273	**0.312**	0.324	0.357	0.388
27	0.224	0.269	**0.308**	0.320	0.353	0.384
28	0.220	0.266	**0.305**	0.316	0.349	0.380
29	0.218	0.263	**0.301**	0.312	0.345	0.376
30	0.215	0.260	**0.298**	0.309	0.341	0.372

[a] In this and the other accompanying tables, the newly generated or corrected values are indicated in boldface. [b] Sample size.

Table II. Critical Values of Dixon's r_{11} Parameter As Applied to a Two-Tailed Test at Various Confidence Levels, Including the 95% Confidence Level

	confidence level					
N[a]	80% (α = 0.20)	90% (α = 0.10)	95% (α = 0.05)	96% (α = 0.04)	98% (α = 0.02)	99% (α = 0.01)
4	0.910	0.955	**0.977**	0.981	0.991	0.995
5	0.728	0.807	**0.863**	0.876	0.916	0.937
6	0.609	0.689	**0.748**	0.763	0.805	0.839
7	0.530	0.610	**0.673**	0.689	0.740	0.782
8	0.479	0.554	**0.615**	0.631	0.683	0.725
9	0.441	0.512	**0.570**	0.587	0.635	0.677
10	0.409	0.477	**0.534**	0.551	0.597	0.639
11	0.385	0.450	**0.505**	0.521	0.566	0.606
12	0.367	0.428	**0.481**	0.498	0.541	0.580
13	0.350	0.410	**0.461**	0.477	0.520	0.558
14	0.336	0.395	**0.445**	0.460	0.502	0.539
15	0.323	0.381	**0.430**	0.445	0.486	0.522
16	0.313	0.369	**0.417**	0.432	0.472	0.508
17	0.303	0.359	**0.406**	0.420	0.460	0.495
18	0.295	0.349	**0.396**	0.410	0.449	0.484
19	0.288	0.341	**0.386**	0.400	0.439	0.473
20	0.282	0.334	**0.379**	0.392	0.430	0.464
21	0.276	0.327	**0.371**	0.384	0.421	0.455
22	0.270	0.320	**0.364**	0.377	0.414	0.446
23	0.265	0.314	**0.357**	0.371	0.407	0.439
24	0.260	0.309	**0.352**	0.365	0.400	0.432
25	0.255	0.304	**0.346**	0.359	0.394	0.426
26	0.250	0.299	**0.341**	0.354	0.389	0.420
27	0.246	0.295	**0.337**	0.349	0.383	0.414
28	0.243	0.291	**0.332**	0.344	0.378	0.409
29	0.239	0.287	**0.328**	0.340	0.374	0.404
30	0.236	0.283	**0.324**	0.336	0.369	0.399

[a] Sample size.

a one-tailed test; but, for very small sample sizes, the effect *may* be slightly less than double *depending upon the distribution pattern of the data* since the tail containing the perceived deviant is determined by the sample rather than by independent knowledge of the true distribution of the population. This argument implies that, at worst, assuming a doubling of the α risk upon going from a one-tailed test to a two-tailed test will result in critical values that are slightly

Table III. Critical Values of Dixon's r_{12} Parameter As Applied to a Two-Tailed Test at Various Confidence Levels, Including the 95% Confidence Level

	confidence level					
N^a	80% (α = 0.20)	90% (α = 0.10)	95% (α = 0.05)	96% (α = 0.04)	98% (α = 0.02)	99% (α = 0.01)
5	0.919	0.960	**0.980**	0.984	0.992	0.996
6	0.745	0.824	**0.878**	0.891	0.925	0.951
7	0.636	0.712	**0.773**	0.791	0.836	0.875
8	0.557	0.632	**0.692**	0.708	0.760	0.797
9	0.504	0.580	**0.639**	0.656	0.702	0.739
10	**0.464**[b]	0.537	**0.594**	0.610	0.655	0.694
11	0.431	0.502	**0.559**	0.575	0.619	0.658
12	0.406	0.473	**0.529**	0.546	0.590	0.629
13	0.387	0.451	**0.505**	0.521	**0.564**[c]	**0.602**[c]
14	0.369	0.432	**0.485**	0.501	0.542	0.580
15	0.354	0.416	**0.467**	0.482	0.523	0.560
16	0.341	0.401	**0.452**	0.467	0.508	0.544
17	0.330	0.388	**0.438**	0.453	0.493	0.529
18	0.320	0.377	**0.426**	0.440	0.480	0.516
19	0.311	0.367	**0.415**	0.429	0.469	0.504
20	0.303	0.358	**0.405**	0.419	0.458	0.493
21	0.296	0.349	**0.396**	0.410	0.449	0.483
22	0.290	0.342	**0.388**	0.402	0.440	0.474
23	0.284	0.336	**0.381**	0.394	0.432	0.465
24	0.278	0.330	**0.374**	0.387	0.423	0.457
25	0.273	0.324	**0.368**	0.381	0.417	0.450
26	0.268	0.319	**0.362**	0.375	0.411	0.443
27	0.263	0.314	**0.357**	0.370	0.405	0.437
28	0.259	0.309	**0.352**	0.365	0.399	0.431
29	0.255	0.305	**0.347**	0.360	0.394	0.426
30	0.251	0.301	**0.343**	0.355	0.389	0.420

[a] Sample size. [b] In Dixon's original table (*13*), the r_{12} critical value at the 80% confidence level for n = 10 is 0.454. However, the cubic regression curve based on the 40%, 60%, 90%, 96%, 98%, and 99% confidence level critical values for n = 10 as well as a regression curve fitted to the 80% cricital values versus sample size indicates that this value should be 0.464. Therefore, it is concluded that the 80% critical value originally published for r_{12} at n = 10 contains a typographical error. The same r_{12} critical value for 95% confidence is obtained either by using the corrected 80% value or by omitting it altogether from the regression analysis. [c] The r_{12} critical values in Dixon's original table (*13*) for n = 13 at the 98% and 99% confidence levels are 0.554 and 0.612, respectively. However, cubic regression curves fitted to the critical values at these two confidence levels as a function of sample size yield corrected values of 0.564 and 0.602, respectively, indicating that the original values contained typographical errors. These latter values were used in resolving the r_{12} critical value for the 95% confidence level at n = 13 (although the same 95% value was obtained by omitting these values and including the critical values for the 40% and 60% confidence levels).

too high, and the resulting decisions that are made will be overly conservative. As shown by comparison of Dixon's original one-tailed critical values for 95% confidence (*15*) with his two-tailed critical values for the 90% confidence level (*16b*), it is clear that Dixon assumed a doubling of the α risk.]

Interpolation of Critical Values at the 95% Confidence Level. As noted above, Dixon was able to obtain exact solutions for the various critical values only for the cases where n = 3 or 4. Although the general form of the equation for $n \geq 5$ has been presented (*15*), the specific expressions for the central density function vary with each sample size, and these expressions have not been published. Therefore, in the current work, appropriate two-tailed critical values of Q at the 95% confidence level were initially estimated by plotting the (two-tailed) critical values for the 99%, 98%, 96%, 90%, and 80% confidence levels as generated by Dixon. The graphically interpolated values for 95% Q were then checked by using regression analysis to determine the best fitting empirical polynomial functions to the values that Dixon published; these functions were then solved for the appropriate critical values of Q at the 95% confidence level.

With the exception of cases in which the original Dixon tables contained apparent errors (vide infra), it was found that a cubic function provided an optimal fit to the Q values in this region. Interestingly, the use of a quadratic function generally yielded the same values of 95% Q, to three significant figures, as those obtained from a cubic function despite a notably poorer fit; fourth-power functions also produced the same 95% values to within ±0.001 but were less sensitive to errors in the tabular data. It was also noted that the cubic functions could generally be extended to include the data for the 60% and 40% confidence levels but, except as noted below, these values were not included in fitting the cubic regression curves since the data in the lower confidence level region did not significantly affect the calculation of the 95% values.

Based on the foregoing analysis, the critical values of Q at the 95% confidence level were then calculated from the cubic regression curve for each sample size and were found to be within ±0.001 of the values obtained graphically. Cubic functions were subsequently fitted to the 99%, 98%, 96%, 90%, and 80% confidence level data for each of the other ratio functions defined by Dixon. To check the veracity of the generated equations, the critical values for 96% and 90% were also calculated and, in each case (except as noted below), were found to be within ±0.001 of Dixon's tabular values.

In using this approach, it was noted that cubic equations could not be fitted to the r_{20} critical values for $n \geq 19$. After examining Dixon's tabular data carefully, it was discovered that the 90% confidence level critical values showed a discontinuity in this region. To circumvent this problem, the 60% confidence level values were included in fitting the r_{20} data to cubic equations. In the region $4 \leq n \leq 18$, the critical values of r_{20} at the 95% confidence level obtained by including and excluding the 60% critical values were within ±0.0002, i.e., undetectable to three significant figures (see Figure 1). For $n \geq 19$, the 90% values were then omitted. In this manner, cubic expressions were generated that provided excellent fits to the data and permitted *both* the 95% and new 90% critical values to be computed. Interestingly, in comparing the newly

Table IV. Critical Values of Dixon's r_{20} Parameter As Applied to a Two-Tailed Test at Various Confidence Levels, Including the 95% Confidence Level

	confidence level					
N[a]	80% (α = 0.20)	90% (α = 0.10)	**95% (α = 0.05)**	96% (α = 0.04)	98% (α = 0.02)	99% (α = 0.01)
4	0.935	0.967	**0.983**	0.987	0.992	0.996
5	0.782	0.845	**0.890**	0.901	0.929	0.950
6	0.670	0.736	**0.786**	0.800	0.836	0.865
7	0.596	0.661	**0.716**	0.732	0.778	0.814
8	0.545	0.607	**0.657**	0.670	0.710	0.746
9	0.505	0.565	**0.614**	0.627	0.667	0.700
10	0.474	0.531	**0.579**	0.592	0.632	0.664
11	0.449	0.504	**0.551**	0.564	0.603	0.627
12	0.429	0.481	**0.527**	0.540	0.579	0.612
13	0.411	0.461	**0.506**	0.520	0.557	0.590
14	0.395	0.445	**0.489**	0.502	0.538	0.571
15	0.382	0.430	**0.473**	0.486	0.522	0.554
16	0.370	0.418	**0.460**	0.472	0.508	0.539
17	0.359	0.406	**0.447**	0.460	0.495	0.526
18	0.350	0.397	**0.437**	0.449	0.484	0.514
19	0.341	**0.387**[b] (0.379)	**0.427**[b]	0.439	0.473	0.503
20	0.333	**0.378**[b] (0.372)	**0.418**[b]	0.430	0.464	0.494
21	0.326	**0.371**[b] (0.365)	**0.410**[b]	0.422	0.455	0.485
22	0.320	**0.364**[b] (0.358)	**0.402**[b]	0.414	0.447	0.477
23	0.314	**0.358**[b] (0.352)	**0.395**[b]	0.407	0.440	0.469
24	0.309	**0.352**[b] (0.347)	**0.390**[b]	0.401	0.434	0.462
25	0.304	**0.346**[b] (0.343)	**0.383**[b]	0.395	0.428	0.456
26	0.300	**0.342**[b] (0.338)	**0.379**[b]	0.390	0.422	0.450
27	0.296	**0.338**[b] (0.334)	**0.374**[b]	0.385	0.417	0.444
28	0.292	**0.333**[b] (0.330)	**0.370**[b]	0.381	0.412	0.439
29	0.288	**0.329**[b] (0.326)	**0.365**[b]	0.376	0.407	0.434
30	0.285	**0.326**[b] (0.322)	**0.361**[b]	0.372	0.402	0.428

[a] Sample size. [b] Starting with n = 19, the r_{20} critical values for both the 90% and 95% confidence levels were calculated from the cubic regression curves fitted to the critical values published by Dixon *(13)* corresponding to the two-tailed 60%, 80%, 96%, 98%, and 99% confidence levels (but omitting the published 90% confidence values). For the 90% confidence level, the values originally published by Dixon are indicated in parentheses underneath the newly generated values. From a comparison of the two sets of values, it is obvious that the critical values in the original table were shifted up one row in the column corresponding to the two-tailed 90% confidence level (see text).

generated 90% confidence level values of r_{20} with the original tabular data, the results reveal that, in Dixon's original table *(15)*, the two-tailed 90% confidence level values (corresponding to Dixon's one-tailed α = 0.05) for 19 ≤ n ≤ 30 were accidentally displaced upward by one row (see Table IV).

In any set of data in which the critical values at the 90% and/or 96% confidence levels, as calculated from the cubic regression equation, differed by more than ±0.001 from Dixon's tabular values, both the cubic equation and the original tabular data were carefully checked for error. In this manner, five additional typographical errors were uncovered in Dixon's original tables. Corrected values were obtained in two ways: (i) by generating a new cubic equation for the specific sample size omitting the suspected tabular value (while extending the cubic fit to include the 60%—and, in some cases the 40%—confidence level value); and (ii) by generating a cubic or quartic equation to fit the critical values at the specific confidence level as a function of sample size. Identical values were obtained with both approaches and, in all five cases, the corrected values revealed that one digit had been incorrectly typeset in the original paper *(15)*. The corrected values are indicated in Tables III and VI.

The resultant critical values of Q (r_{10}) at the 95% confidence

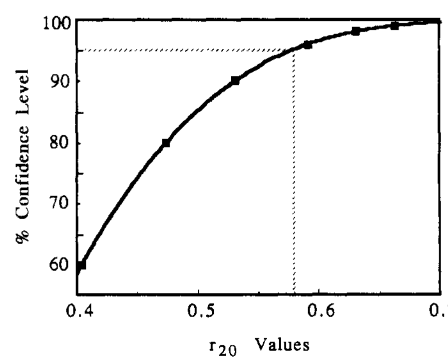

Figure 1. Typical plot of critical values as a function of the confidence level. The curve shown is the regression curve for r_{20} at n = 10 (including the values for 60%, 80%, 90%, 96%, 98%, and 99%) corresponding to the cubic equation: % confidence level = −343.7575 + 1835.0349r_{20} − 2548.5722r_{20}^2 + 1188.477r_{20}^3.

level and similarly generated critical values for the r_{11}, r_{12}, r_{20}, and r_{22} functions are included in Tables I–VI along with the values published by Dixon corresponding to the 99%, 98%, 96%, 90%, and 80% confidence levels—*all confidence levels shown being applicable to a two-tailed test.* All values have

Table V. Critical Values of Dixon's r_{21} Parameter As Applied to a Two-Tailed Test at Various Confidence Levels, Including the 95% Confidence Level

	confidence level					
N^a	80% (α = 0.20)	90% (α = 0.10)	95% (α = 0.05)	96% (α = 0.04)	98% (α = 0.02)	99% (α = 0.01)
5	0.952	0.976	**0.987**	0.990	0.995	0.998
6	0.821	0.872	**0.913**	0.924	0.951	0.970
7	0.725	0.780	**0.828**	0.842	0.885	0.919
8	0.650	0.710	**0.763**	0.780	0.829	0.868
9	0.594	0.657	**0.710**	0.725	0.776	0.816
10	0.551	0.612	**0.664**	0.678	0.726	0.760
11	0.517	0.576	**0.625**	0.638	0.679	0.713
12	0.490	0.546	**0.592**	0.605	0.642	0.675
13	0.467	0.521	**0.565**	0.578	0.615	0.649
14	0.448	0.501	**0.544**	0.556	0.593	0.627
15	0.431	0.483	**0.525**	0.537	0.574	0.607
16	0.416	0.467	**0.509**	0.521	0.557	0.580
17	0.403	0.453	**0.495**	0.507	0.542	0.573
18	0.391	0.440	**0.482**	0.494	0.529	0.559
19	0.380	0.428	**0.469**	0.482	0.517	0.547
20	0.371	0.419	**0.460**	0.472	0.506	0.536
21	0.363	0.410	**0.450**	0.462	0.496	0.526
22	0.356	0.402	**0.441**	0.453	0.487	0.517
23	0.349	0.395	**0.434**	0.445	0.479	0.509
24	0.343	0.388	**0.427**	0.438	0.471	0.501
25	0.337	0.382	**0.420**	0.431	0.464	0.493
26	0.331	0.376	**0.414**	0.424	0.457	0.486
27	0.325	0.370	**0.407**	0.418	0.450	0.479
28	0.320	0.365	**0.402**	0.412	0.444	0.472
29	0.316	0.360	**0.396**	0.406	0.438	0.466
30	0.312	0.355	**0.391**	0.401	0.433	0.460

[a] Sample size.

Table VI. Critical Values of Dixon's r_{22} Parameter As Applied to a Two-Tailed Test at Various Confidence Levels, Including the 95% Confidence Level

	confidence level					
N^a	80% (α = 0.20)	90% (α = 0.10)	95% (α = 0.05)	96% (α = 0.04)	98% (α = 0.02)	99% (α = 0.01)
6	0.965	0.983	**0.990**	0.992	0.995	0.998
7	0.850	0.881	**0.909**	0.919	0.945	0.970
8	0.745	0.803	**0.846**	0.857	0.890	0.922
9	0.676	0.737	**0.787**	0.800	0.840	0.873
10	0.620	0.682	**0.734**	0.749	0.791	0.826
11	0.578	0.637	**0.688**	0.703	0.745	0.781
12	0.543	0.600	**0.648**	0.661	0.704	0.740
13	0.515	0.570	**0.616**	0.628	0.670	0.705
14	0.492	0.546	**0.590**	0.602	0.641	0.674
15	0.472	0.525	**0.568**	0.579	0.616	0.647
16	0.454	0.507	**0.548**	0.559	0.595	0.624
17	0.438	0.490	**0.531**	0.542	0.577	0.605
18	0.424	0.475	**0.516**	0.527	0.561	0.589
19	0.412	0.462	**0.503**	0.514	0.547	0.575
20	0.401	0.450	**0.491**	0.502	0.535	0.562
21	0.391	0.440	**0.480**	0.491	0.524	0.551
22	0.382	0.430	**0.470**	0.481	0.514	0.541
23	0.374	0.421	**0.461**	0.472	0.505	0.532
24	0.367	0.413	**0.452**	**0.464**[b]	0.497	0.524
25	0.360	0.406	**0.445**	0.457	0.489	0.516
26	0.354	0.399	**0.438**	0.450	**0.482**[b]	0.508
27	0.348	0.393	**0.432**	0.443	0.475	0.501
28	0.342	0.387	**0.426**	0.437	0.469	0.495
29	0.337	0.381	**0.419**	0.431	0.463	0.489
30	0.332	0.376	**0.414**	0.425	0.457	0.483

[a] Sample size. [b] The r_{22} critical values listed in Dixon's original table (*13*) for n = 24 at the 96% confidence level (two-tailed) and for n = 26 at the 98% confidence level are 0.484 and 0.486, respectively. The values shown in this table (0.464 and 0.482, respectively) were obtained by a cubic regression curve fitted to the critical values at these two confidence levels as a function of sample size. These same values are also generated from the cubic regression lines fitted to the critical values for each of these sample sizes omitting the questionable values and including the values at the 40% and 60% confidence levels. It is concluded that the original tabular critical values for r_{22} for these two sample sizes and confidence levels were the result of typographical errors.

been carefully cross-checked, and the 95% critical values are judged to be accurate within ±0.001 relative to the accuracy of the values at the other confidence levels (which Dixon stated (*15*) were themselves generally accurate to within ±0.001). The values shown cover the entire range of sample sizes ($3 \le n \le 30$) included in Dixon's original article.

It is suggested that the critical values of Q and the related r criteria which have been generated for the 95% confidence level should be used routinely by practicing analytical chemists in testing for the rejection of outliers since this confidence level provides a reasonable compromise between ultraconservatism and the overzealous rejection of deviant values. These values should also be incorporated into future analytical chemistry treatises and textbooks dealing with tests for the rejection of data to provide a uniform set of critical values at this standard confidence level.

As a concluding comment, it should be noted that recent studies on the variance of the arithmetic mean after rejection of outliers suggest the superiority of two more recently proposed criteria (Huber-type skipped mean and Shapiro-Wilk rules) for rejection decisions (*19*), particularly for larger samples containing multiple outliers. Nonetheless, the simplicity of Dixon's range ratio tests argues strongly for their continued use in many analytical applications.

ACKNOWLEDGMENT

I express my appreciation to my colleagues, George Schenk and David Coleman, for helpful suggestions regarding this manuscript.

LITERATURE CITED

(1) (a) Miller, J. C.; Miller, J. N. *Statistics for Analytical Chemistry*, 2nd ed.; Wiley: New York, 1988; pp 62-64, 218. (b) Anderson, R. L. *Practical Statistics for Analytical Chemists*; Van Nostrand Reinhold: New York, 1987; pp 31-32. (c) Caulcutt, R.; Boddy, R. *Statistics for Analytical Chemists*; Chapman and Hall: London, 1983; pp 66-67, 248.
(2) Barnett, V.; Lewis, T. *Outliers in Statistical Data*, 2nd ed.; Wiley: New York, 1984.
(3) (a) Skoog, D. A.; West, D. M.; Holler, F. J. *Analytical Chemistry: An Introduction*, 5th ed.; Saunders: Philadelphia, 1990; p 56. (b) Skoog, D. A.; West, D. M.; Holler, F. J. *Fundamentals of Analytical Chemistry*, 5th ed.; Saunders: New York, 1988; pp 13-16. (c) Hargis, L. G. *Analytical Chemistry*; Prentice-Hall: Englewood Cliffs, NJ, 1988; p 56. (d) Fritz, J. S.; Schenk, G. H. *Quantitative Analytical Chemistry*, 5th ed.; Allyn & Bacon: Boston, 1987; pp 45-46. (e) Potts, L. W. *Quantitative Analysis: Theory and Practice*; Harper & Row: New York, 1987; pp 78-80. (f) Rubinson, K. A. *Chemical Analysis*; Little, Brown: Boston, 1987; pp 162-164. (g) Day, R. A., Jr.; Underwood, A. L. *Quantitative Analysis*, 5th ed.; Prentice-Hall: Englewood Cliffs, NJ, 1986; pp 29-31. (h) Manahan, S. E. *Quantitative Chemical Analysis*; Brooks Cole: Monterey, CA, 1986; pp 74-75. (i) Kennedy, J. H. *Analytical Chemistry: Principles*, 2nd ed.; Harcourt, Brace, Jovanovich: New York, 1990; pp 35-39. (j) Harris, D. C. *Quantitative Chemical Analysis*; Freeman: San Francisco, 1982; pp 51-52. (k) Ramette, R. W. *Chemical Equilibrium and Analysis*; Addison-Wesley: Reading, MA, 1981; pp 53-54. (l) Christian, G. D. *Analytical Chemistry*; Wiley: New York, 1980; pp 78-79. (m) Flaschka, H. A.; Barnard, A. J., Jr.; Sturrock, P. E. *Quantitative Analytical Chemistry*; Willard Grant: Boston, 1980; pp 19-20.
(4) Dixon, W. J. *Ann. Math. Stat.* **1950**, *21*, 488-506
(5) King, E. P. *J. Am. Stat. Assoc.* **1953**, *48*, 531-533.
(6) Deming, W. E. *Statistical Adjustment of Data*; Wiley: New York, 1943 (republished by Dover: New York, 1964); p 171.
(7) Parratt, L. G. *Probability and Experimental Errors in Science*; Wiley: New York, 1961; pp 176-178.
(8) Hawkins, D. M. *Identification of Outliers*; Chapman and Hall: London, 1980.
(9) Beckman, R. J.; Cook, R. D. *Technometrics* **1983**, *25*, 119-149.
(10) (a) Burr, I. W. *Applied Statistical Methods*; Academic: New York, 1974; pp 194-195. (b) Hamilton, W. C. *Statistics in Physical Science*; Ronald: New York, 1964; pp 45-49.
(11) Dixon, W. J. In *Contributions to Order Statistics*; Sarhan, A. E., Greenberg, B. G., Eds.; Wiley: New York, 1962; pp 299-342 (see pp 314-317).
(12) Natrella, M. G. *Experimental Statistics*; National Bureau of Standards Handbook 91; NBS: Washington, DC, 1963; Chapter 17.
(13) "Student". [Gossett, W. S.] *Biometrika* **1908**, *6*, 1.
(14) Taylor, J. R. *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*; University Science: Mill Valley, CA, 1982; pp 142-145.
(15) Dixon, W. J. *Ann. Math. Stat.* **1951**, *22*, 68-78.
(16) (a) Dixon, W. J. *Biometrics* **1953**, *9*, 74-89. (b) Dean, R. B.; Dixon, W. J. *Anal. Chem.* **1951**, *23*, 636-638.
(17) Peters, D. G.; Hayes, J. M.; Hieftje, G. M. *Chemical Separations and Measurements: Theory and Practice of Analytical Chemistry*; Saunders: Philadelphia, 1974; p 36.
(18) Youmans, H. L. *Statistics for Chemistry*; Charles E. Merrill: Columbus, OH, 1973; p 65 ff.
(19) Hampel, F. R. *Technometrics* **1985**, *27*, 95-107.

RECEIVED for review June 15, 1990. Accepted October 16, 1990.

Frequency-Domain Spectroscopic Study of the Effect of *n*-Propanol on the Internal Viscosity of Sodium Dodecyl Sulfate Micelles

M. J. Wirth,* S.-H. Chou, and D. A. Piasecki

Department of Chemistry and Biochemistry, University of Delaware, Newark, Delaware 19716

The rotational diffusion behavior of tetracene in sodium dodecyl sulfate micelles is studied as a function of the *n*-propanol content of the micellar solution. Fluorescence anisotropy measurements using frequency-domain spectroscopy show that tetracene reorients faster with increasing concentration of *n*-propanol. This result is consistent with micellar liquid chromatographic studies and lends insight into the role of *n*-propanol as a mobile-phase modifier. The components of the rotational diffusion tensor are determined from the double-exponential anisotropy decay. These confirm the validity of the Debye-Stokes-Einstein model and allow calculation of the viscosity of the micelle interior. The viscosity decreases from 8 to 4 cP as the concentration of *n*-propanol increases from 0 to 10% (v/v). The components of the rotational diffusion tensor indicate that the solvation environment of tetracene in the micelle is structurally disordered.

*To whom correspondence should be addressed.

INTRODUCTION

The importance of micelles in analytical chemistry has burgeoned both in spectroscopy and in separation science. Micelles are used to enhance fluorescence (*1*), thermal lensing (*2*), and room-temperature phosphorescence (*3*). In liquid chromatography, micelles modify the organic content of the mobile phase (*4*), allowing more rapid gradient elution (*5*). Micelles can serve as the pseudostationary phase in electrokinetic chromatography (*6*), providing selectivity for separation of polar organic solutes. The nature of solute interactions with micelles is thus a vital area of research.

An important question entails the dynamics of solutes interacting with micelles. One of the drawbacks of micellar liquid chromatography has been poor column efficiency due to slow mass transfer between the surfactant-modified stationary phase and the mobile phase (*7*). Dorsey et al. demonstrated that the addition of at least 1% of *n*-propanol gives a 40% increase in the number of theoretical plates for the solute benzene in micellar liquid chromatography (*8*). They