
METHODS FOR DETERMINING THE GOLD CONTENT OF JEWELRY METALS

By Meredith E. Mercer

Verifying the gold content of karat-gold jewelry is a problem for the entire industry. This study compares four widely available testing methods: density (measured by hydrostatic weighing), chemical reactivity (touchstone testing), capacitive decay (Tri-Electronics G-XL-18 Gold Tester), and X-ray fluorescence (Seiko Instruments SEA 2001 Jewelry Assay System). The latter three methods all proved useful within certain limitations; hydrostatic weighing was least suitable. None of these methods reliably determined gold content within the 3-ppt tolerance allowed by U.S. plumb laws.

In jewelry, the integrity of the metal is as important as the identity and integrity of the gemstones (figure 1). The misrepresentation of gold content ("underkarating") has been a source of concern at all levels of the industry for literally hundreds of years. For the protection of retailers and consumers alike, a testing method that can verify the gold content of finished jewelry at point of sale is needed. Ideally, this method would be rapid, simple, reliable, inexpensive, nondestructive, and widely accepted by the jewelry industry. The study reported here was conducted to determine whether existing methods could meet these requirements. Four testing methods currently used in the trade—density, chemical reactivity, capacitive decay, and X-ray fluorescence—were applied to the same set of well-characterized reference standards and the results compared.

Gold-testing methods such as fire assay, touchstone testing, and hydrostatic weighing have been used since antiquity. Tests for plated material include cutting, heating, and ringing (Oddy, 1983; Revere, 1990). Much work has been done on the development of analyses to detect very small quantities of gold for mining (Young, 1980). Methods of analytical chemistry (such as X-ray transmission, Compton scattering, neutron activation analysis, inelastic scattering, atomic absorption, and particle-induced X-ray emission) can be quite accurate. However, these methods usually call for complex and expensive equipment, as well as destructive sample preparation (de Jesus, 1985; Kahn et al., 1981; Demortier, 1984). To the author's knowledge, no work evaluating techniques commonly used on jewelry metals in tests of carefully characterized karat gold alloys has been previously published.

ABOUT THE AUTHOR

Ms. Mercer is a research associate in the Research Department of the Gemological Institute of America, Santa Monica, California.

Acknowledgments: S. Armstrong of S. Alexis Manufacturing helped gather data. H. Chitrik of Chitra Trading Corp. kindly donated funds used for this research. Leach and Garner Technology provided reference standards and analyses. The Worshipful Company of Goldsmiths' London Assay Office provided analyses and performed some of the testing. Tri-Electronics Corp. assisted with one of the tests.

A technical paper on this research was presented at the Santa Fe Symposium on Jewelry Manufacturing Technology, May 1992, and will be available in the proceedings of this symposium.

Gems & Gemology, Vol. 28, No. 4, pp. 222-233.

© 1993 Gemological Institute of America

BACKGROUND

Underkarating is the practice of knowingly or unknowingly misrepresenting the gold content of a finished piece to be higher than it actually is. A Jewelers Vigilance Committee (JVC) study found that as much as 50%–70% of the untrade-

Figure 1. Karat weight is an important consideration in gold jewelry, for customer and retailer alike. There are also different alloy colors of gold. Illustrated here, clockwise from the upper right, are a 14k green gold bee-shaped pin set with pavé diamonds; an 18k green gold pin set with diamonds and a black opal; a 14k yellow gold chain; a 14k ring of red, white, and yellow gold; and a 24k gold Chinese tael. The "bee" pin, chain, and ring are courtesy of The Gold Rush, Northridge, California; the opal-set pin is by Silverhorn, Santa Barbara, California. Photo © GIA and Tino Hammid.



marked jewelry they tested was underkarated (i.e., contained less gold than its karat mark stated; Tolhurst, 1986). In contrast, JVC reported in the same article that they found no underkarating in a random sample of trademarked items that they checked. Court cases involving underkarating are currently pending in San Francisco and Honolulu ("Police beat," 1989; Badham, 1990). The fact that there may be only slight differences in visual appearance between golds of greatly different karat grades underscores the difficulty of detecting underkarating in daily commerce (figure 2).

United States law recognizes alloys containing at least 417 parts per thousand (ppt; 41.7% or 10k) gold as karat gold, and allows a tolerance of 3 ppt on

unsoldered items (see Box A for an explanation of the relationship between parts-per-thousand, percent, and karat designations). A 7-ppt tolerance is allowed on items containing solder. Thus, a 14k (nominally, 583-ppt) gold item that has been made or repaired with solder may be as low as 576-ppt gold. Where a quality mark stating gold content appears, it must be accompanied by a manufacturer's trademark (Jewelers Vigilance Committee, 1987; Shor, 1988). Although this trademark has been legally required since 1906, it is largely unenforced.

The oldest institution charged with protecting the consumer from underkarating is the Worshipful Company of Goldsmiths in London. In 1300, King Edward I of England required that all gold and silver

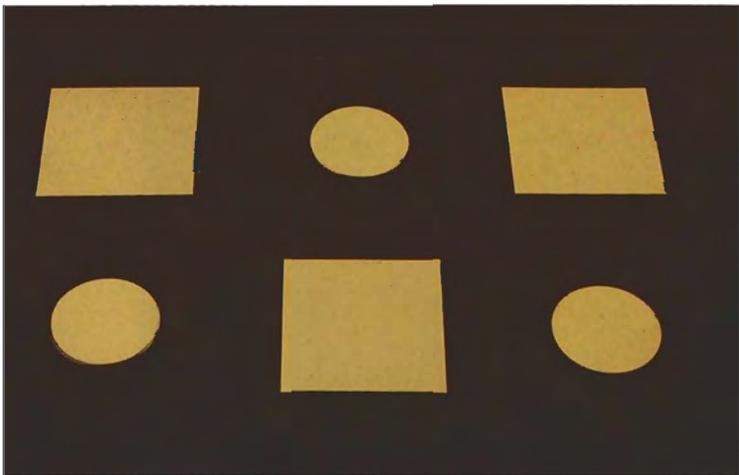


Figure 2. Although these karat-gold pieces are similar in color and appearance, they are actually 10k, 14k, and 18k (left to right). Gold courtesy of David H. Fell and Company, Inc., Los Angeles, California; photo by Maha Smith-DeMaggio.

wares be tested and marked by the goldsmith's guild, a precursor of the Worshipful Company (Johnson, 1980). Later, wares had to be sent to the guild hall for testing and marking (the probable origin of the English word *hallmark*; Hare, 1985).

Today, every item of gold, silver, or platinum sold in the United Kingdom must first be sent to one of four assay offices for hallmarking. Those that pass the assay are marked with four (or sometimes five) marks: (1) a maker's or sponsor's mark; (2) in certain cases, a symbol indicating the fineness of the item; (3) the fineness in parts per thousand; (4) the symbol of the office where testing was done; and (5) a letter in a particular typeface that indicates the year of marking. Not only does this information protect consumers, but it is also very valuable to jewelry historians. Unlike U.S. law, British law allows no negative tolerance (e.g., an item that is 749-ppt gold cannot be marked 18k, which is nominally 750 ppt; Assay Offices of Great Britain, 1988).

BOX A: THE NOMENCLATURE OF GOLD CONTENT

The gold content (also known as "fineness") of karat gold is described in several ways:

Percent (%) is simply the ratio of gold (by weight) to the total metal present, expressed as parts of 100. Thus, an alloy that is three parts gold and one part alloy metal is:

$$3/(3+1) = 3/4 = 75/100 = 75\% \text{ (by weight)}$$

Parts per thousand (ppt) is similar to percent, but compares the total metal (by weight) present to 1000, rather than 100, parts. This increases the expressed precision. Thus, an alloy containing seven parts fine gold and five parts alloy metal is:

$$7/(7+5) = 7/12 = 0.583 = 583 \text{ ppt (by weight)}$$

This could also be expressed as 58.3%. Legal tolerances and requirements are usually expressed in ppt.

Karat (k, also abbreviated kt) is perhaps the most widely known measure of gold content—as well as the least straightforward to calculate and envision. It is also based on the ratio of fine gold to total metal present (by weight), but it divides the total metal present into 24 parts. Thus, each karat is 1/24 (by weight) of the whole (4.17%, or 41.7 ppt). The 583-ppt gold alloy (seven parts gold and five

parts alloy metal) discussed above can be described in karats as:

$$7/(7+5) = 7/12 = 14/24 = 14k$$

This measure of gold content is spelled "karat" in the United States to help avoid confusion with "carat," the measure of weight commonly used for gemstones. In Great Britain and British Commonwealth countries, it is spelled "carat" for both applications.

Table A-1 correlates these three measures of gold content.

TABLE A-1. Comparison of expressions of gold content.

Karat (k)	Percent (wt. %)	Parts per thousand (ppt)
8	33.3	333
9	37.5	375
10	41.7	417
12	50	500
14	58.3	583
18	75	750
22	91.7	917
24	100 ^a	1000 ^a

^aNote: Most countries allow gold that is 99% (990 ppt) pure to be described as "24 karat," or "pure gold."



Figure 3. Ten karat-gold reference standards were used for this study (right to left): 8k yellow, 10k white, 10k yellow, 14k green, 14k red, 14k white, 14k yellow, 18k yellow, 22k yellow, and 24k. Photo by Maha Smith-DeMaggio.

International laws regarding karat standards vary widely. Seven European countries signed the 1972 "International Convention on the Control of Marking of Articles of Precious Metals," which established standards and authorized laboratories within member countries to apply a convention mark certifying gold content. This mark allows each member country to accept goods from other member countries without further verification. Today, hallmarking is compulsory in France, Ireland, Portugal, Spain, Holland, and Great Britain (P.V.A. Johnson, pers. comm., 1992); some other European countries have voluntary hallmarking programs. The effect of European economic unification on hallmarking laws is still being worked out (Johnson, 1992).

MATERIALS AND METHODS

Four methods for testing the gold content of karat golds were selected for this study: density, touchstone testing, capacitive decay, and X-ray fluorescence spectrometry. The criteria used in choosing these four methods were nondestructiveness, ease of operation, and, except for X-ray fluorescence, low cost and portability. The same 10 reference standards were tested by each method.

The reference standards in this study are karat golds commonly used in jewelry. They were custom cast, rolled, and given a brushed finish for GIA at Leach and Garner Technology, North Attleboro, Massachusetts. The standards range from 8k to 24k, and include yellow, white, red, and green golds (figure 3).

The composition of each reference standard was determined at Leach and Garner by fire assay and

directly coupled plasma spectroscopy, and then confirmed by fire assay and atomic-absorption spectroscopy at Goldsmith's Hall, London. The compositions found are shown in table 1.

Fire assay is one of the oldest methods known for testing gold content. The earliest written mention of the technique dates from 1380 B.C. (Wälchli and Vuillemier, 1987). In fire assay, the metal is oxidized at high temperature (cupelled) to separate base metals from noble metals. Any silver present is then removed from the residual metal by dissolution in nitric acid. The weight of the remaining fine gold is compared to the weight of the original sample to establish the metal's original gold content.

In atomic-absorption (AA) spectroscopy, the sample is dissolved in a liquid solution and then vaporized in a flame. The resulting gas absorbs light in proportion to its elemental composition, which can be measured to produce a quantitative chemical analysis.

TABLE 1. Compositions of karat-gold reference standards used for this study^a.

Reference standard	Composition (wt.%)				
	Au	Ag	Cu	Zn	Ni
24k yellow	99.99 ± 0.1 (99.99)	— ^b	—	—	—
22k yellow	91.55 ± 0.2 (91.63)	2.85 ± 0.5 (3.07)	5.10 ± 0.5 (4.90)	0.50 ± 0.2 (0.50)	—
18k yellow	75.00 ± 0.2 (75.03)	6.75 ± 0.5 (6.67)	15.35 ± 0.5 (15.34)	2.90 ± 0.5 (2.92)	—
14k yellow	58.50 ± 0.2 (58.58)	3.95 ± 0.5 (3.74)	31.50 ± 0.5 (31.69)	6.05 ± 0.5 (5.94)	—
14k white	58.40 ± 0.2 (58.39)	—	22.70 ± 0.5 (22.66)	7.65 ± 0.5 (7.95)	11.25 ± 0.5 (11.00)
14k red	58.40 ± 0.2 (58.42)	—	40.80 ± 0.5 (40.87)	0.80 ± 0.3 (0.76)	—
14k green	58.40 ± 0.2 (58.42)	32.20 ± 0.5 (32.28)	9.20 ± 0.5 (9.17)	0.20 ± 0.1 (0.22)	—
10k yellow	41.70 ± 0.2 (41.72)	5.80 ± 0.5 (5.62)	44.00 ± 0.5 (44.21)	8.50 ± 0.5 (8.42)	—
10k white	41.60 ± 0.2 (41.63)	—	33.45 ± 0.5 (32.66)	8.45 ± 0.5 (8.80)	16.50 (16.93)
8k yellow	33.40 ± 0.2 (33.40)	6.90 ± 0.5 (6.95)	49.95 ± 0.5 (49.87)	9.75 ± 0.5 (9.78)	—

^aReference standards cast and analyzed (fire assay and DC plasma) by Leach and Garner Technology, North Attleboro, MA; uncertainty in measured value (±) varies with alloy content and method of measurement. Values in parentheses are from the Worshipful Company of Goldsmiths Assay Office, London, UK; analysis by fire assay and atomic absorption.

^b— not detected.

For gold, silver, copper, tin, and cadmium, AA has very high sensitivity (approximately 1 atomic part per million [ppm—"parts per million atoms," not to be confused with ppt]; Veillon and Skogerboe, 1987). The disadvantages of AA include the fact that it can analyze for only one element at a time, and the possibility that error could be introduced when the sample is diluted to the low concentration necessary (Willard et al., 1981, p. 141).

Directly coupled plasma (DCP) spectroscopy, which is the technique used by the metallurgical laboratory that prepared the standards for this study, is similar to atomic absorption. As with AA, the sample is placed in a liquid solution, but it is then excited by high-temperature plasma—a gas ionized by highly concentrated electromagnetic energy—rather than by light. This excitation causes the emission of light (i.e., photons) with wavelengths that are characteristic of the sample's composition. The advantages of DCP are that it can analyze several elements at once and that it can handle solutions in high concentrations (up to 25%), which increases the accuracy that can be obtained for jewelry metals. Like AA, it is sensitive to approximately 1 ppm for most metals commonly found in jewelry. The primary disadvantage of DCP is its unsuitability to automated operation (Willard et al., 1981, p. 163). A related method, inductively coupled plasma (ICP) spectroscopy, is also frequently used, with similar results.

These three analytical methods, while more accurate and repeatable than the four methods examined in this study, are also much more complex and expensive. Therefore, they are not suitable for general use by most gemologists.

TESTS AND RESULTS

This study was conducted to examine the performance of each of the four testing methods on a set of well-characterized reference standards. Only the 10 reference standards were tested. No metals other than karat gold, no plated items or finished jewelry, were examined. Other shapes and finishes would probably produce somewhat different results.

Each method tested has unique advantages and limitations, and direct comparison between methods is difficult. The purpose of this work is not a competitive comparison of methods, but an individual evaluation of each.

Density. Common alloying elements (such as nickel, copper, zinc, and silver) have significantly lower den-

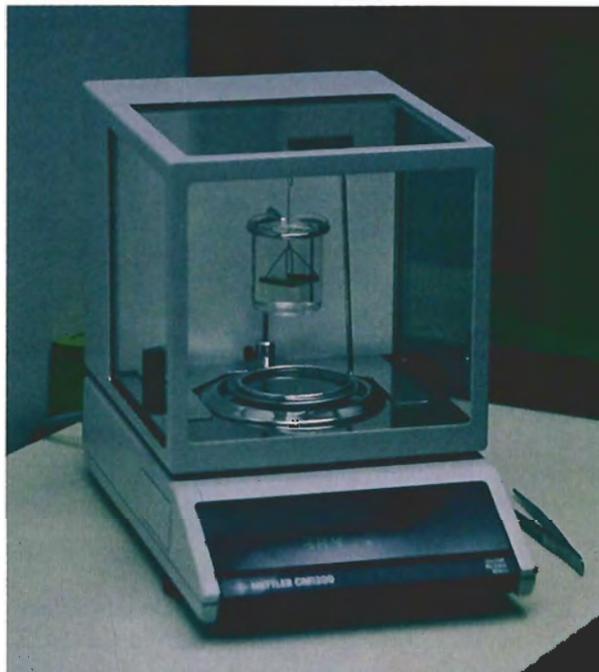


Figure 4. A Mettler balance with hydrostatic-weighing accessories was used to test the specific gravities of the karat-gold reference samples. Photo by Maha Smith-DeMaggio.

sities than gold. Density is often expressed as specific gravity (S.G.), a numerical comparison of the density of a substance to that of water. Hydrostatic weighing works on the principle that density can be measured by the buoyant force acting on an object immersed in a fluid.

According to legend, the ancient Greek scientist Archimedes (c. 287–212 B.C.) discovered this principle in response to a gold-testing problem (van den Waerden and Heath, 1983). King Heiron II of Syracuse asked Archimedes if there was a way to determine, without damaging the item, whether his new crown was made of the fine gold he had specified or of the adulterated alloy he suspected had been used.

Archimedes puzzled over the problem one day while visiting the public baths. Stepping into a tub, he watched water run out over the top—and realized that an immersed object displaces fluid in a way dependent only on its volume, and irrespective of its weight. Thus, he could find the volume of the crown by immersing it in a vessel filled completely with water, measuring the overflow, and comparing the volume of the overflow to the volume of an amount of pure gold that weighed precisely the same as the crown. If

the crown were pure gold, the two volumes would be the same; but if the crown were gold alloyed with a lighter metal, such as silver or copper, the same weight of alloy must have a greater volume than pure gold. Because gold is denser, that is, has a higher specific gravity, it displaces less water than an equal weight of silver or copper.

This idea struck Archimedes so strongly that he leaped up and, forgetting his clothing, ran home through the streets shouting "Eureka, Eureka!" ("I have found it!").

Method. The reader is referred to any of a number of books on gemological testing for description of the technique of hydrostatic weighing (e.g., Webster, 1983; Liddicoat, 1990) and to Sinkankas (1986) for an in-depth historical review of the use of this technique in gemology. Weighings for this study were performed on the equipment shown in figure 4.

Each standard was weighed in air and water four times (except the 14k green and 8k yellow samples, which were weighed only three times each). Densities were calculated according to the following equation:

$$\text{density} = \frac{\text{weight in air}}{\text{weight in air} - \text{weight in water}}$$

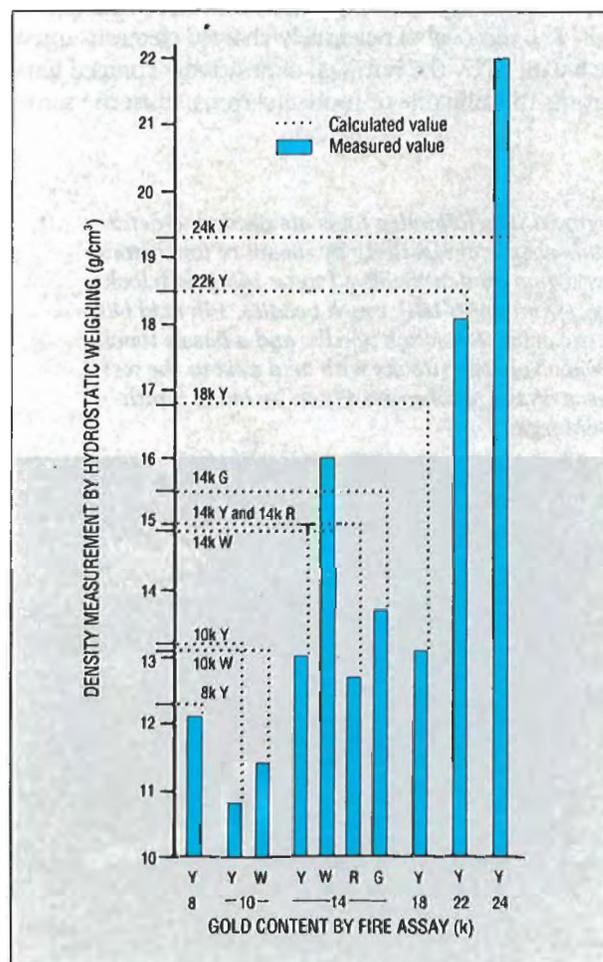
The means and ranges for the densities determined for each standard are shown in figure 5. In addition, density was calculated for each reference standard according to its known composition (as determined by fire assay, DCP, and AA spectroscopy) and is shown for comparison in figure 5. Calculated density is based on the measured concentration of each element in an alloy, multiplied by the known density of that element.

Results. As figure 5 shows, most of the mean measured data do not match the calculated densities (dashed lines). In eight of 10 cases, the hydrostatic results indicate less gold than is actually present. The 3-ppt tolerance allowed by U.S. plumb law is too small to show on this chart. For example, the 6-ppt (± 3 ppt) plumb range for the 14k yellow sample corresponds to a density range from 14.96 g/cm³ to 14.88 g/cm³. In contrast, the mean measured value is 13.0 g/cm³.

Evaluation. Although theoretically straightforward, hydrostatic weighing presents practical difficulties. The balance pan must be weighed first, and its weight subtracted from the final water weight. Air bubbles clinging to the test piece will falsely decrease weight, while any water droplets on the wires holding the

weighing tray will falsely increase weight. The accuracy of the density method could be improved by making a greater number of measurements per sample, but the time required to do so would make this method impractical for most jewelry applications. Moreover, only objects without stones can be tested. This method also cannot detect gold-plated items filled with a metal (such as lead) that has a density similar to that of gold.

Figure 5. Densities determined by hydrostatic weighing are compared to densities calculated from the analytically measured contents of the 10 reference standards. The mean result of four weighings is shown for each sample. Y, W, R, G indicate yellow, white, red, and green golds, respectively. Note that the densities calculated apply only to the karat-gold samples used for this study; different karat golds of the same color and gold content may have different densities.



Chemical Reactivity. The property that sets gold apart from most other metals is its relative inertness to chemical reaction, which makes it highly resistant to corrosion. This property can be used to test jewelry metals, because the gold in the alloy will react much more slowly when exposed to acid than will the other metals. In addition, the more alloying metal present in proportion to gold, the greater will be the magnitude of the reaction to the acid. This type of testing is commonly known as “touchstone testing.”

Corrosion is a property that occurs only in metals. Thus, to understand why and how corrosion occurs, one must first understand that metals as a group are unique among chemical elements. Their electrons are not rigidly bound in place, but are free to move about. In addition, electrons in different elements have different intrinsic energy levels, like different natural bodies of water at different altitudes. For example, when two different metals are placed in electrical contact, electrons flow from the metal with the higher energy level (e.g., copper) to the one with the lower energy level (e.g., gold), as water flows downhill. The removal of negatively charged electrons must be balanced by the removal of positively charged ions (atoms missing one or more electrons) from the same

piece of metal. The changes in color and texture of metals that we commonly associate with corrosion result from this loss of material. (It is this flow of electrons that also produces current in simple batteries, and it is this eventual depletion of material that causes batteries to “go dead”). The higher the energy level of a given metal, the more it will react in this way. Thus, we may judge these energy levels by observing chemical reactivity.

Touchstone testing for gold content, which exploits this chemical reactivity in metals, is the origin of the phrase *acid test*. Touchstone testing is done by visually comparing the reactions of a test metal and the reactions of metals of known gold content to the same acids simultaneously. References to the comparison of uncorroded metal streaks can be found as early as 600 B.C. (Wälchli and Vuillemier, 1985). The use of acid to measure the chemical reactivity of the streaks dates to the 14th century. The acid reacts with the atoms of other metals (such as silver, copper, or zinc) in a gold alloy by creating the flow of electrons and ions described above between microscopic regions richer in these other metals and regions richer in gold. The resulting corrosion can be observed as a discoloration.

Figure 6. The following tools are needed to determine chemical reactivity by means of touchstone testing on an item believed to be 14k gold (clockwise from upper left): touch needles, 14k acid (43% nitric acid), 14k touch needle, and a basalt touchstone. Note the streaks with acid next to the test piece on the touchstone. Photo by Maha Smith-DeMaggio.



Method. In touchstone testing, the test piece is drawn carefully across the surface of a fine-grained black “touchstone” (commonly, a piece of basalt; figure 6), producing a streak on the stone. The streak is bracketed by streaks from “touch needles”—strips of karat gold of known gold content. A drop of acid is drawn across all the streaks, and their subsequent reaction is carefully watched. A judgment is then made as to which touch-needle reaction most closely matches that of the test piece. The test is repeated, with different acids or touch needles, as necessary. Nitric acid, hydrochloric acid, and aqua regia are the most commonly used acids. It is reported that accuracy to at least 10–20 ppt, and (with experience) as close as 5 ppt, can be made with this method, given the availability of touch needles that are graduated to this degree of fineness (Wälchli and Vuillemier, 1985).

Tests were performed by the author on the stone shown in figure 6 with 9k, 14k, 18k, and 22k yellow and red-gold touch needles manufactured by Bergeon (No. 6675-9). The acids were mixed according to the formulations used at Goldsmith’s Hall. They included 9k (24% aqueous HNO_3), 14k (43% aqueous HNO_3), 18k (43% aqueous HNO_3 with 0.06% NaCl), and 22k (43% aqueous HNO_3 with 1% NaCl). Note that these acids, while used in very small amounts, are

hazardous materials that must be handled properly.

Each sample was tested four times by drawing it across the stone to create an even, consistent metal streak. On one side, a similar streak was drawn from a touch needle of possibly equal or lower karat weight. (An educated guess is the best estimate initially.) On the other side, a streak was drawn from the next-higher-karat touch needle. A drop of the acid estimated to match the sample karatage was placed on the touchstone near the streaks. From this drop, a thin, even line of acid was pulled (with the acid dropper) across the sample and touch-needle streaks. At least two comparison touch-needle streaks and two different acid reactions were used to test each sample.

Reactions were observed as they took place, and the appearance of the various streaks was noted after the acid had been soaked up with filter paper, in combined fluorescent/incandescent lighting. Between tests, as needed, the touchstone was cleaned with 600-grit silicon carbide sandpaper lubricated with olive oil.

Goldsmith's Hall also tested all but the 18k sample (which was not available at the time). Four different operators performed one test each on the nine other samples, using acids mixed to the same formulas described above.

Results. The results of the author's tests and those done at Goldsmith's Hall are shown in figure 7. All of the data from tests performed at GIA lie within a statistically acceptable range of the known values, as do most of the results from Goldsmith's Hall. The exceptions are the data for the 10k yellow and 10k white samples obtained by Goldsmith's Hall. The 14k red samples and white samples showed a wide range of values, as did the 18k, 22k, and 24k (yellow) samples. As was the case with hydrostatic weighing, all of the ranges are outside the 3-ppt tolerance specified by U.S. plumb law.

Evaluation. Tests performed by a trained but inexperienced technician (the author) tended to be closer to assayed values than those obtained by highly skilled and experienced workers at Goldsmith's Hall. This is probably due to the difficulty of creating a truly blind test for the author, and indicates one bias inherent in this kind of testing. Especially notable are results for the two 10k alloys, which were almost unanimously reported as 9k in the Goldsmith's Hall trials. Ten-karat gold is almost unknown in Britain, where the

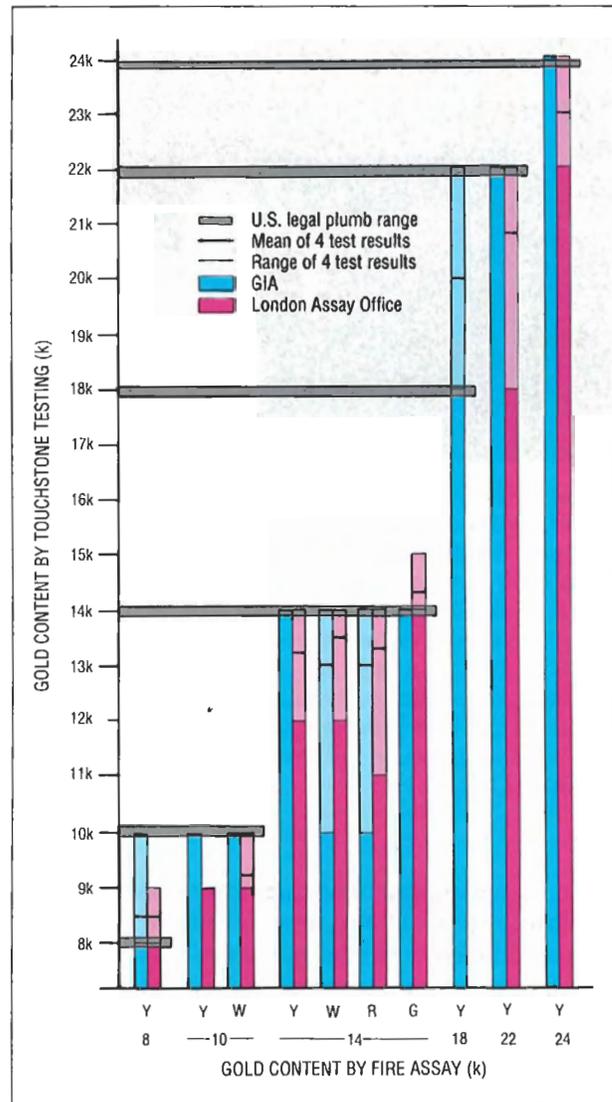


Figure 7. Gold content was measured by touchstone testing at GIA and at the London Assay Office of Goldsmith's Hall. The mean and range of four tests for each sample are shown (the 18k sample was not available for testing at the Assay Office). The horizontal gray bars indicate the 6-ppt (± 3 ppt) range allowed by U.S. plumb laws.

most common karat gold is 9k (which, as noted earlier, is not legally defined as karat gold in the U.S.).

Because touchstone results depend heavily on the skill, experience, and bias of the observer, they tend to be subjective. Also, some precision is lost with high-karat golds because of the increasing similarity of the reactions. Again, note that the acids are hazardous materials that must be used with caution.



Figure 8. The components of the G-XL-18 Gold Tester, used for capacitive decay testing, include (from left to right): the G-XL-18 tester unit, an electrical hook-up with test piece and electrolyte gel, and a gel dispenser/probe wand. For this test, the operator positions the test piece in the clip and applies a drop of gel. With the probe touching the gel, the operator then presses the button that corresponds to the color of the test piece. The results are displayed on the tester unit. Photo by Maha Smith-DeMaggio.

Even so, measurement of chemical reactivity by the touchstone test can be relatively reliable if the operator is sufficiently skilled and free of bias.

An important advantage of this method is that it can be used to test jewelry of almost any shape, size, and finish, including stone-set pieces. In addition, plating can be detected by skillful observation of the inhomogeneous reaction pattern that occurs within the metal streak.

Capacitive Decay. This method was invented by Leonid Radomyshelsky, a Jewish emigré from the Soviet Union who had been trained as an electrical engineer. After Mr. Radomyshelsky opened a jewelry store in the U.S., he became concerned about verifying the gold content of the jewelry he sold. Using his engineering expertise, he devised an electronic method to measure the "nobility" of metals.

Method. The G-XL-18 Gold Tester (figure 8), made by Tri-Electronics Corp. of San Diego, California, operates by measuring the capacitive decay rate of the test piece. When a test piece is placed in electrical contact with the platinum cathode, the tester

builds up a type of electrical charge known as "capacitance" between an electrolyte gel and the cathode. This capacitance decays at a measurable rate that varies with the gold content of the test piece. The G-XL-18 tester measures the decay rate, compares it to an internal computer file, and displays the karatage that matches the measured decay rate most closely.

Two sets of tests were run on the Gold Tester, both on the same instrument. For the first set, performed by the author, nine trials were run on each sample. For the second set, performed by Mr. Radomyshelsky himself, five to seven trials (the number he recommends for this instrument) were run on each test piece.

Because the G-XL-18 is designed to test gold only in the range of 10k to 18k, we did not include results for samples over 18k. However, we did include the 8k sample, because 8k or 9k gold may be encountered both in underkarated 10k goods and in 8k or 9k goods that are legal in various parts of the world. It was understood, however, that while the tester is capable of operating outside its design range, the accuracy of the results might suffer.

Results. The data are presented in chart form in figure 9 (for statistical consistency, only seven of the GIA trials are represented). As might be expected, the means for the 8k sample are off by more than 1k, with a considerable range. Also interesting are the results for the 14k yellow gold, which consistently produced a 12k read-out in the GIA trials. Although all of the results fall within 2k of known values, this is still well outside the 3-ppt tolerance allowed by U.S. law.

Evaluation. The inaccuracy encountered with the 14k yellow sample was traced to a higher copper content in the test piece than in the yellow gold used to calibrate the instrument. When the 14k yellow sample was retested as a red gold, the mean of four tests came to 13.6k, with very little variability. This illustrates the risk inherent in directly comparing the chemistry of one karat-gold piece to another that, although of the same gold content, may have a very different alloy composition. The poor results for the 8k sample simply confirm the manufacturer's stated limitations of the instrument.

One disadvantage of the G-XL-18 is that the presence of plating can only be detected by testing a small area from which the plating has been removed. Also, the curved surfaces of most chains, prong settings, and other small, highly curved items will interfere

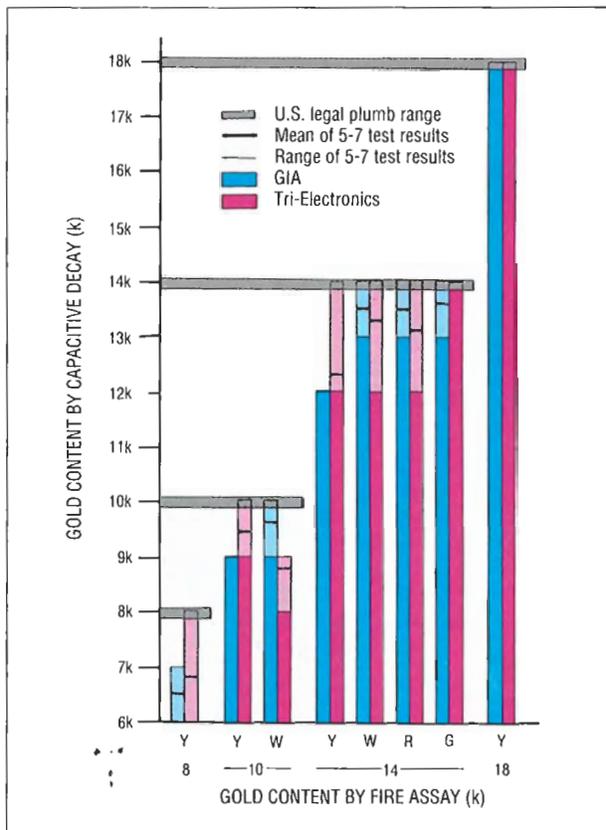


Figure 9. Capacitive decay testing was performed by GIA and Tri-Electronics using the G-XL-18 Gold Tester. The means and ranges of five to seven tests by Tri-Electronics, and seven tests by GIA, are presented in comparison to the 6-ppt (± 3 ppt) range allowed by U.S. plumb law. Although tests were not conducted on samples above the G-XL-18's 10k-to-18k design range, the 8k sample was included to show how the tester responded to a piece that was below the U.S. legal limit for karat gold.

with the capacitive-decay reaction. Tri-Electronics does not recommend the G-XL-18 for testing such items. An advantage of this method is that it can be used to test jewelry of most other shapes and most finishes, as well as stone-set pieces.

X-ray Fluorescence. In 1913, it was discovered that exposure to high-energy radiation can cause atoms to emit X-rays that are characteristic of the source element's atomic number. In 1951, X-ray intensity was correlated to the quantitative chemical composition of the sample. Development in the late 1960s of the lithium-drifted silicon X-ray detector and sophisticated computer systems paved the way for rapid

and relatively accurate quantitative analyses by this method (Goldstein et al., 1981).

Method. When an X-ray strikes an atom, it may cause the atom to emit a secondary X-ray that has an energy unique to the element from which it is emitted. These secondary X-rays can be collected and analyzed for information on sample composition. This is called X-ray fluorescence (XRF) analysis, because atoms in the sample respond to the original, incoming X-rays by fluorescing (emitting) characteristic secondary X-rays.

Tests were performed by the author at Seiko Instruments, Torrance, California, on their SEA 2001 Jewelry Assay System, an XRF system with sophisticated software developed specifically for testing karat gold (see figure 10). One test was run for each reference sample, with the instrument first calibrated to comparable standards when available. Standards were not available for the 24k and 22k yellow, and 14k green, test pieces.

Technicians at Goldsmith's Hall performed a second set of tests (one for each sample, except the 18k yellow) using the same-model instrument, also calibrated to comparable standards.

Figure 10. The SEA 2001 Jewelry Assay System for X-ray fluorescence analysis was the most sophisticated method used in this study. The components of this unit are (clockwise from top right): the liquid-nitrogen reservoir that contains the X-ray detector, the sample chamber (on top of the console), a computer, and a monitor (here displaying a typical spectrum). Photo courtesy of Seiko Instruments Corp.



Results. Figure 11 shows that 80% of the results from tests performed by the author with the Seiko instrument were within 3 ppt of assayed values. The fact that the 24k sample was 11 ppt below, and the 14k green 15 ppt below, their assayed values reflects the lack of calibration standards noted above. The testing at Goldsmith's Hall produced results that fell outside 3 ppt for the 14k green (14 ppt below), the 14k white (5 ppt above), and the 8k yellow (4 ppt below) samples.

Evaluation. The results of XRF analysis are promising, but several caveats must be made. Secondary X-rays travel only very short distances through metal. Therefore, the X-rays available for XRF analysis come from only the first few microns (millionths of a meter) below the sample surface. Thus, only a very small area is tested, which would not represent the bulk of a sample that was plated or otherwise internally inhomogeneous. This surface sensitivity also causes a strong dependence on shape and finish. Results for convex surfaces, unusual finishes, or relief work are generally unreliable.

Moreover, the XRF system must be calibrated to an alloy of similar composition for reliable and accurate quantitative results to be obtained. Seiko Instruments calibrated their tester with standards very similar to the reference standards used in this study. Both alloy sets were obtained from the same supplier. These results, therefore, may not represent practical testing of a variety of alloys from different suppliers. The greater variation in the data from the Goldsmith's Hall tests supports this concern.

The greatest disadvantage of XRF for most in the jewelry industry is that the system is considerably more expensive than the other methods described here, and it requires substantial technical skill both to operate the equipment and to interpret the results.

Despite these limitations, X-ray fluorescence testing can be relatively accurate if the test piece is flat and homogeneous, and the instrument is properly calibrated to a standard of similar shape, finish, and composition.

CONCLUSIONS

From this limited examination of 10 standard reference materials of normal gold-alloy composition ranging from 8k to 24k, the following preliminary conclusions can be drawn:

1. Hydrostatic weighing appears to be unsuitable for karatage determinations, as results are likely to be unreliable.

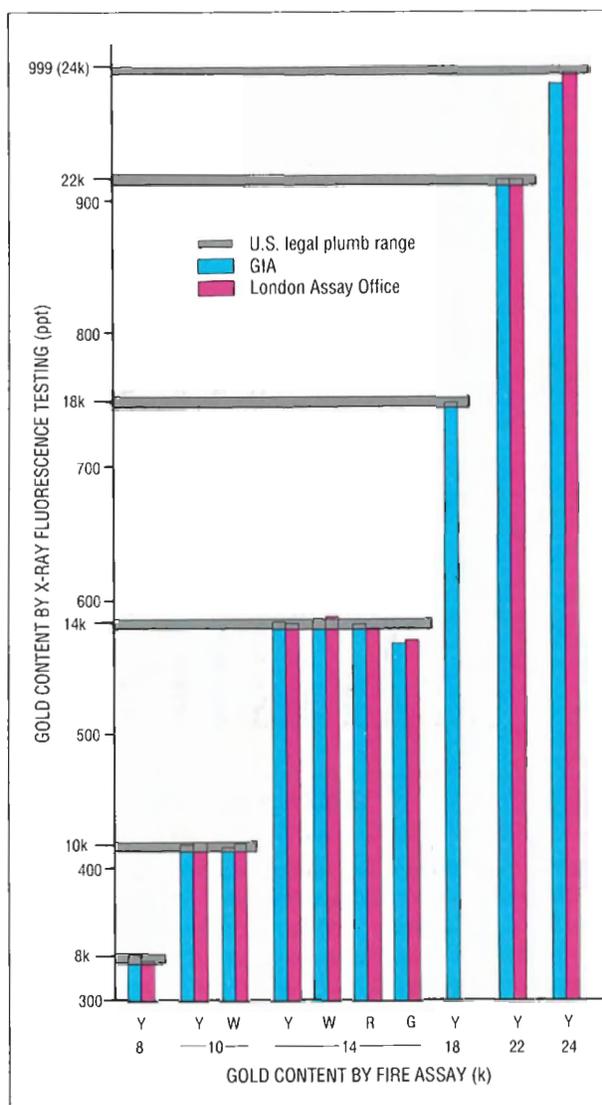


Figure 11. X-ray fluorescence testing was conducted independently by both GIA and the London Assay Office of Goldsmith's Hall. One test of each sample was made (the 18k sample was not available for Assay Office testing). Several, but not all, of the results lie within the gray bar that marks the 6-ppt (± 3 ppt) range allowed by U.S. plumb law.

2. Measurement of chemical reactivity by the touchstone test can be reasonably accurate if the operator is sufficiently skilled and free of bias. Plating can be detected by a skilled and observant operator, and the test is useful on jewelry of most shapes and finishes, as well as on stone-set pieces.
3. Instruments that use the principle of capacitive energy decay give results accurate to 2k in the

range below 18k and above 8k. This method can be used on pieces of many shapes and sizes, and on stone-set pieces, but not on some chains, prong settings, and other small, highly curved items. Plating may not be detected.

4. X-ray fluorescence testing can be accurate to within 15 ppt of gold content if the test piece is flat and homogeneous, and the instrument is calibrated to an alloy of similar composition.

None of the test methods examined here will solve the underkarating problem. In fact, none gave reliable readings that fell completely within the ± 3 -ppt plumb range allowed by U.S. law. However, some of these methods may be of limited use in the retail setting to identify larger discrepancies in gold content.

Most of the errors found in this study were in the form of erroneously low values. Although these would not result in underkarating if they were used by manufacturers and sellers to create and mark their jewelry, buyers might be mistakenly led to believe an item

had been underkarated. Moreover, some values were high, which could lead to unintentional underkarating. Regardless, it is risky to rely on a testing method that produces large errors in any direction.

A great deal more research is needed, both on the evaluation of sophisticated instrumentation that may be able to make the close determinations required by plumb laws, and on the development of methods that are more useful for the retail jeweler. An important limitation of this study is that it does not address actual jewelry in all of its shapes, finishes, and alloy combinations. In addition, no plated metals were tested. Note, however, that a rhodium flash on white gold would not be expected to substantially affect the accuracy of any of these methods, with the possible exception of X-ray fluorescence.

Underkarating is a problem that has existed throughout the history of "precious" metals. It will continue to be a problem until and unless an inexpensive, accurate, reliable, and easily operable testing method is developed.

REFERENCES

- Assay Offices of Great Britain (1988) *Hallmarks on Gold, Silver and Platinum* (pamphlet). The Assay Office, Publications Dept., Goldsmith's Hall, London.
- Badham J. (1990) San Francisco D.A. uses tester in case against IPI gold. *National Jeweler*, Vol. 34, No. 20, pp. 3, 82.
- de Jesus A. (1985) Authentication of gold products by nuclear methods: A feasibility study on gold coins and jewellery. *Gold Bulletin*, Vol. 18, No. 4, pp. 132-139.
- Demortier G. (1984) Analysis of gold jewellery artifacts: Characterization of ancient gold solders by PIXE. *Gold Bulletin*, Vol. 17, No. 1, pp. 27-38.
- Goldstein J., Newbury D., Echlin P., Jog D., Fiori C., Lifshin E. (1981) *Scanning Electron Microscopy and X-Ray Microanalysis*. Plenum Press, New York.
- Hare S.M. (1985) *Goldsmith's Hall in the City of London*. The Worshipful Company of Goldsmiths and Jarrold Colour Publications, Norwich, U.K.
- Jewelers Vigilance Committee (1987) National gold and silver marking act. In *California Jeweler's Association Buying Guide: 1987-88*, pp. 94-96.
- Johnson P.V.A. (1980) Assaying procedures at Goldsmith's Hall for hallmarking purposes. *Proceedings of the Second International Seminar on Sampling and Assaying of Precious Metals II, March 18-19, 1980*. San Francisco, IPMI (International Precious Metals Institute), Allentown, PA, pp. 147-157.
- Johnson P.V.A. (1992) Hallmarking and analysis of finished jewelry. *Precious Metals 1992: Proceedings of the 16th International Precious Metals Conference*, IPMI (International Precious Metals Institute), Allentown, PA pp.205-211.
- Kahn H., Dulude G., Conley M., Sotera J. (1981) Determination of metals at high concentration by atomic absorption spectroscopy. *Canadian Research*, Vol. 14, No. 1, pp. 27-31.
- Liddicoat R.T. Jr. (1990) *Handbook of Gem Identification*, 12th ed. rev. Gemological Institute of America, Santa Monica, CA.
- Oddy A. (1983) Assaying in antiquity. *Gold Bulletin*, Vol. 16, No. 2, pp. 52-59.
- Police beat: under-karated gold (1989). *Honolulu Advertiser*, April 16, 1989, p. A13.
- Revere A. (1990) Testing precious metals in the work shop. *Jewelers' Circular-Keystone*, Vol. 161, No. 8, pp. 180-186.
- Shor R. (1988) Subkarat gold: a "time bomb" set to explode. *Jewelers' Circular-Keystone*, Vol. 154, No. 9, pp. 54-56.
- Sinkankas J. (1986) Contributions to a history of gemology: Specific gravity—origins and development of the hydrostatic method. *Gems & Gemology*, Vol. 22, No. 3, pp. 156-165.
- Tolhurst M.E. (1986) Rash of underkarating threatens trade: suppliers. *National Jeweler*, Vol. 30, No. 16, pp. 1, 31, 34-36.
- van den Waerden B.L., Heath T.L. (1983) Archimedes. In *Encyclopedia Britannica*, 15th ed., Macropedia, Vol. 1, pp. 1087-1088.
- Veillon C., Skogerboe R. (1987) Atomic spectroscopy. In S. Parker, Ed., *Spectroscopy Source Book*, McGraw-Hill Book Co., New York.
- Wälchli W., Vuillemeir P. (1985) Touchstone testing of precious metals. *Aurum*, No. 24, pp. 36-45.
- Wälchli W., Vuillemeir P. (1987) Assaying gold by cupellation. *Aurum*, No. 29, pp. 56-64.
- Webster R. (1983) *Gems: Their Sources, Descriptions and Identification*, 4th ed. Rev. by B.W. Anderson, Butterworths, London, pp. 650-652.
- Willard H., Merritt L., Dean J., Settle F. (1981) *Instrumental Methods of Analysis*. D. Van Nostrand, New York, pp. 127-174.
- Young R. (1980) Analysis for gold: A review of methods. *Gold Bulletin*, Vol. 13, No. 1, pp. 9-14.