COMPARISON OF VISUAL GRADING METHODS

Brian F. Chabot and Stephen L. Childs
Cornell University

Visual determination of maple syrup color is the most common method for grading syrup. It is the only grading method approved by the USDA and by several states. Standardized color grading began in 1910 based on caramel solutions and the glass standards came into use in 1950. Over the years a variety of grading kits have been introduced. The most common kit in use is the Vermont Temporary kit. One using plastic color filters has replaced the USDA glass kit. Lovibond is still available in glass and a new glass kit is being manufactured by White Mountain Maple Products (Rimouski, Quebec).

There does not appear to have been a published comparison of the different grading kits subsequent to the research that introduced caramel standards followed by glass standards. Consequently, we undertook a comparison of most of the grading kits that are currently in use. These included USDA glass and plastic, Vermont Temporary, Grimm glass (for 2 oz bottles), Berliner glass, Lovibond (US and Canadian) and a White Mountain Maple Products glass kit.

METHODS

Two series of tests were conducted. In the first, we obtained 84 samples from New York; 33 directly from producers and 51 as coded samples collected by NYS Agriculture & Markets for lead testing. We used five different visual grading kits, three Hanna analyzers, and two laboratory spectrophotometers to classify the color density of each syrup. Results from the comparison of visual and spectrophotometric methods were published previously (Chabot and Childs, 2006). For the visual comparisons, we used two observers and different lighting conditions, including colored light from a computer-driven LCD projector. Syrup samples were placed in one-ounce bottles, which fit in the sample holders of most of the kits, except that we used the 3 cm cuvette for the US version of the Lovibond kit.

In the second series of tests, we used eight kits, including both the US and Canadian versions of the Lovibond kit. We used some of the same syrup samples as in the first test, but we obtained additional syrup samples from producers. In addition, we constructed samples by blending syrups. Samples were blended in order to have a greater number of samples close to the grade boundaries to create a greater challenge for any differences that might exist between the kits. We also used the type of sample container for which the kits were originally designed as we suspected that some of the differences between kits seen in the first series were due to inappropriate sample containers. New 2-oz bottles are too large for the older kits, which is why the 1-oz bottle was used in the first series of tests. We obtained 3-cm glass cuvettes designed for Berliner kits that fit in all the older kits designed for 2-oz bottles.

RESULTS

In the first series of tests, the biggest difference was between the grade on the container and the grades we determined with any of the visual grading kits (Table 1). All samples in Table 1 were graded against a clear bright northern sky by the same person. Seventy-five percent of samples were graded darker than the producer label, 20 samples were the same grade, and one sample was lighter. Samples obtained directly from producers and stored under refrigeration had slightly more grade differences than did those samples obtained from Agriculture & Markets. A few samples were more than 2 grades darker.

Table 1. Number of samples in four color grades using producer grade and five grading kits.

Grade	Producer	USDA	Vermont	Lovibond	Berliner	Grimm
Light	23	13	11	8	20	13
Medium	45	18	18	21	27	31
Dark	16	27	30	27	17	19
B	0	26	25	28	20	21

There were some differences among the grading kits in the first series (Table 1). The Lovibond kit graded somewhat darker and the Berliner kit somewhat lighter than the other kits. Only the Vermont and Lovibond kits used sample

containers for which these kits were designed. We used a 1-oz container in the other kits, which has a shorter light path-length than the 2-oz containers, leading to lighter-appearing samples. The relative performance of the kits in these tests must be interpreted with this in mind.

Comparing the Vermont and Lovibond kits, grades were different in 18 cases; 21% of samples. There were three differences in the lighter grades; in each case Lovibond produced a darker grade. The other 15 differences were in the medium or darker grades and here the Lovibond produced lighter grades. We also compared two USDA glass kits where 5 (6%) differences in grade determination occurred, which was the fewest in any of the comparisons. Three differences were in the dark/extra dark grades. One kit produced darker grades in four instances.

We were able to rank the samples using the spectrophotometric measurements of light transmission. We also visually ranked the samples from lightest to darkest. Overall samples that were visually darker had lower transmissivity (%T), but there were many cases where the two rankings did not match. For example, a sample with a %T of 68% appeared visually darker than two samples with %T of 66.3 and 64.5.

The same person working with the same kit can obtain slightly different results in somewhat different bright sky conditions (Table 2). In this test, 22 samples (26%) were graded differently. Differences occurred at all grades. Although it seems obvious that the samples that differed must be close to a grade boundary, in only 3 cases was grade determination noted to be difficult.

Table 2. Number of samples in four color grades using the Vermont Temporary Kit on different days in different locations.

Grade	Vermont- campus sky	Vermont- Arnot sky
Light	11	11
Medium	18	12
Dark	30	32
B	25	29

Two individuals can grade differently under the same lighting conditions with the same kit (Table 3). There were 11 (13%) differences in grade determination. Steve produced a lighter grade in 8 cases and a darker grade in 3 cases. Five of these were judged to be difficult determinations by one or both graders. There were 7 other difficult determinations where both graders produced the same grade.

Table 3. Number of samples in four color grades for two people using the Vermont Temporary Kit with the same lighting conditions.

Grade	Vermont (Brian)	Vermont (Steve)
Light	12	11
Medium	12	18
Dark	34	30
B	26	25

Background light color can affect grade determination (Tables 4, 5). The USDA and Vermont kits were used for these tests. The bottom rows in the tables show the number of cases where the grade changed compared to skylight as the reference. With the USDA kit, 10 to 32% of the samples changed grade depending upon background light. White light produced roughly the same number of differences at all three color boundaries. Light blue and red had more differences at the medium/dark boundary than at the other boundaries. The other colors had most of the changes at the dark/extra dark boundary. The Vermont kit produced a somewhat larger number of differences, which might be expected because there is an additional grade to distinguish. There were two differences at the light/medium transition. Most of the differences were medium/dark or darker grades. More samples were placed in the darker grades with these artificial light sources.

Table 4. Number of samples in four color grades using a USDA glass kit and nine background light colors.

Grade	Sky	White	Purple	MBlue	LBlue	Green	Yellow	Red	Brown
Light	13	8	12	11	8	12	12	11	11
Medium	18	23	17	19	42	16	18	13	16
Dark	27	29	23	21	6	31	35	31	37
B	26	24	32	33	28	25	19	29	20
changes	0	15	9	12	27	11	13	12	14

Table 5. Number of samples in five color grades using a Vermont Temporary Kit and nine background colors.

Grade	Sky	White	Purple	MBlue	LBlue	Green	Yellow	Red	Brown
Light	11	11	11	11	11	11	12	11	11
Medium	18	9	9	15	13	10	8	10	10
Dark	27	36	30	29	32	34	36	35	35
B	12	15	18	15	16	16	15	16	15
Substd	16	13	16	14	12	13	13	12	13
Changes	0	20	23	14	19	19	21	21	21

At the recommendation of Dr. Ellis Loew, we tested grading using a blue filter between our eyes and the samples to focus our attention on light amount rather than shifts in syrup color hue. The reduction in the amount of light reaching our eyes made grading more difficult.

In the second series of tests, we included additional grading kits. We also used the Berliner glass sample cuvette for all the kits designed for 3-cm light paths. All samples were graded against three light sources: open bright sky, and a 100 watt incandescent bulb, a color-corrected fluorescent bulb and a 23-watt compact fluorescent bulb all behind a white plastic diffuser. In addition to these light sources, the Lovibond samples were also evaluated with a white light source from Lovibond designed for color determinations with their kits. The percent transmittance of all samples were determined with a Hanna Analyzer.

The kits in these tests gave more similar results than in the first tests (Table 6). This is indicated by the number of differences from the USDA glass kit as shown in the bottom row. Because a higher proportion of samples were made to be at a boundary between two grades and the difficulty of determining grade was noted, the differences shown are a conservative estimate. For example we didn't count it as a difference if one kit showed a sample to be on the light/medium boundary and another kit graded the syrup as light.

In general, samples seen as close calls with one kit were also close calls with other kits. The differences reflect those cases where the decision was not close, and these were few in number. There were no differences between any of the kits in determining which samples were light amber. There were two samples at the medium/dark boundary and one at the dark/B boundary that caused most of the discrepancies.

Table 6. Number of samples in four color grades using eight kits: White Mountain Maple (WMM), Lovibond Canadian (LovC), Vermont Temporary (Vt), Berliner (Berlin), USDA glass (USglas), USDA plastic (USpls), Lovibond USDA with fluorescent light (LovUf) and with Lovibond light unit (LovUl), and Grimm.

Grade	WMM	LovC	Vt	Berlin	USglas	USpls	LovUf	LovUl	Grimm
Light	5	5	5	5	5	5	5	5	5
Med	8	10	9	13	12	7	11	10	13
Dark	20	14	14	14	13	18	14	15	12
B	7	11	12	8	10	10	10	10	10
Diff	2	2	1	2	0	2	0	0	0

Several different light sources were used in this research. Lovibond, color-corrected fluorescent, and compact fluorescent (23 watt) behind a white plastic panel gave results similar to that of clear sunny sky. Incandescent light (100 watt) was yellower in color and gave greater differences in grade determination.

DISCUSSION

Visual grading is not an exact method for classifying syrup color density. Color is based on how our brains interpret signals from four different light-absorbing pigments in our eyes. Human visual pigments detect light over a broad spectrum and are sensitive to both amounts of light and changes in the proportion of different wavelengths. Because our eyes differ in color sensitivity and our brains differ in how we interpret colors, two people can come to different conclusions about a sample grade. The color of background lighting can affect the results and there are slight differences in the color absorption of caramel, glass, and plastic standards. Additionally, the color hue varies between syrups and this can interact differently with different kits.

In the first series of tests we found some differences between the kits that we thought might be significant. However, these differences are most likely due to using incorrect glass containers for the samples. A second series of tests using

correct containers gave smaller and, we judge, not significant differences between grading kits. So using the correct sample container is important.

The original USDA caramel and glass standards were developed using sample containers with a 3-cm light path through the syrup. Lovibond (US) and the Berliner containers are 3 cm exactly. The USDA and other kit manufacturers adopted the 2 oz bottle as a close approximation of the more expensive 3-cm cuvette. Although the internal path length for light may be the same in the newer 2-oz bottles, the walls are thicker and we were unable to locate any current source of bottles that fit in the older kits. The new USDA kit is designed for the larger bottles.

We have found some producers also using the 1-oz bottles in kits designed for 2-oz bottles for the same reason we used them. The 1-oz bottle has a shorter light path length through the syrup causing the syrup to be graded lighter. There are Grimm kits designed around a 1-oz bottle that we did not test.

Lighting conditions are another factor of considerable importance in syrup grading. The original USDA glass color standards were developed using a National Bureau of Standards lighting source that is not commercially available. Lovibond sells a light source designed to produce a balanced white light. This light source is less bright and grade decisions are less easily made. It also is very expensive. Given the variation in grading decisions we found with various

light sources, it may be preferable to use an artificial light source that is the same each time a grading decision is made. Of the light sources we used, a compact fluorescent lamp behind a white plastic diffuser was the best in terms of ease of use and cost.

The significant number of samples that were graded darker than the producer label merits further study. Causes may include a combination of oxidation in the container, optimistic grading of samples close to the boundary, and grading while the syrup is hot, which makes the syrup less dense. A better understanding of these grade differences seems necessary.

The kits differed in usability. The easiest kit to assess grade is the Lovibond, because the sample is placed next to the standard visually without any break and the sample is optically uniform. The most difficult to use is the new USDA kit, because there are wide black panels between the sample and the standard and the bottles are optically imperfect, which creates varying light density across the samples and standards. The sample and all the standards are relatively close to each other in space in the Berliner kit and does not require the sample cuvette to be moved to compare it with the standards. However, the sample is much larger than the standards and the greater amount of light from the sample made some comparisons difficult. The Vermont kit has the advantage of low purchase price and easy availability of inexpensive sample bottles. Comparison of kits of different ages suggests that the supplier's advice to use a new kit each year is probably correct, which removes the initial cost advantage.

The bottom line on the performance of different grading kits is that when used with the proper sample containers and good lighting all of the visual kits produced close to the same results.

ACKNOWLEDGEMENT

We thank Prof. Ellis Loew for assistance in understanding the effect of light color on grading decisions, and Dr. Kathleen Rasmussen for editing the manuscript.

REFERENCE

Chabot, B. F. and S. Childs. 2006. Relating spectrophotometer readings to visual grading of maple syrup. Maple Syrup Digest 18A(2):34-37.

NEW EDITION OF MAPLE MANUAL AVAILABLE

The new edition of the North American Maple Syrup Producers Manual is now available. Printed in full color, the new manual is 329 pages long, contains 13 chapters and 5 appendices and over 150 illustrations. This edition of the Manual is entirely new, containing contributions from twenty educators, researchers and maple producers from across the maple region.

The Manual may be purchased from your association, many dealers and equipment manufacturers.