

The Reliability of Substance and Texture Statements Used in AOS Award Descriptions
By Catherine R. Higgins, PHD
Published in Awards Quarterly, Vol. 38, No. 3, Sept. 2007, pages 217-218

Reliability and validity are familiar concepts to anyone trained in social sciences research. The concept of validity refers to the extent to which a measuring instrument measures the variable it is supposed to measure and measures it accurately. Reliability indicates the accuracy, precision, or consistency of measuring instruments, specifically whether repeated measurements of the same variable produce a measurement with an error rate that falls within a tolerable (and usually narrow) range. (All measurements are assumed to have an error rate.)

For the purposes of orchid award descriptions, reliability is the more relevant concept. The structure of the AOS awards program would suggest that the reliability of descriptions may be a challenge. Judges are volunteers, meaning that they learn measurement techniques informally and are not held financially or professionally accountable for their performance. The judging system is highly decentralized and spread over a large geographic area. Student training takes place within regional centers, all of which develop and implement their own training regimen. Consistency across centers is enforced by regular, informal meetings, usually at orchid shows; a handbook written by a central Judging Committee; a centralized editor to whom all descriptions are submitted, and a quarterly publication in which written descriptions can be informally compared. However, centralized training and retraining of measurement techniques, which is standard practice on social sciences research projects, does not occur.

Three aspects of an award description are: flower size and inflorescence count, color and shape, and substance and texture statements. These could be summarized as: quantitative, qualitative and other. For the quantitative measurements, we have the objective standard of a ruler. Different measurements of segments probably vary by no more than one or two millimeters, assuming, on multifloral plants, that different judges select the same flower to measure. Certainly, it would be useful to know the average margin of error of flower measurements; can we assume, for example, that any measurement is on average accurate to within two millimeters? This is an easy question to test, and it would be an excellent project for student training.

I suspect that flower counts and measurements are also reasonably consistent across measurers. On plants with more than 100 flowers, such as a well-flowered *Dendrobium kingianum* hybrid, counts may differ by several flowers, but these differences could be said to fall within a tolerable range. On these plants, counts are often conducted by several judges, which could, depending on one's perspective, either increase or lower the accuracy — for me, they increase the odds that a count approaches accuracy. Some judges report approximate numbers of flowers on very floriferous plants, usually for a cultural award. In these situations, the count of inflorescences is reported as an exact number, with an approximate number of flowers per inflorescence given. In my opinion,

this approach accurately conveys the cultural achievement and is perfectly appropriate for the circumstances.

For the qualitative measurements, the reliability of color and shape descriptions can be compared against a photographic record. Color-neutral lamps help to neutralize any bias in a lighting source, and spirited discussions around these lamps often result in a color-by-consensus decision. An article in a previous *Awards Quarterly* suggesting a standard method for recording color, using the Pantone® system, made no progress in convincing judges that a better method exists than their subjective assessments.¹ Even when film does lie about color, it usually records shape accurately, providing a corrective, if one is needed, to a verbal description.

Most descriptions conclude with a description of a flower's substance and texture. The *AOS Handbook on Judging and Exhibition* (11th edition) defines substance as “the thickness and firmness of tissue in flowers,” and texture as “the surface qualities of a flower that enhance appearance.” Here, there is no objective standard to measure a description against. Nor is there any standard list of verbal descriptors. Substance is a tactile characteristic, usually assessed by rubbing a flower's segments between two fingers, whereas texture is determined visually, ideally under a lamp.

For some time, I have been curious to know whether substance and texture are consistent for awarded cultivars of the same species or hybrid. We expect size and color to vary, but experience has led me to expect that substance and texture might be consistent. However, I had concerns about the reliability of measurements of this characteristic. For the purposes of this exercise, I hypothesized that substance and texture would be consistent across awarded cultivars, and that the use of terms (i.e., the measurement) would reflect this fact.

To test this question, I tallied the terms used in award descriptions published in *Awards Quarterly* for two hybrids and one species. I selected plants that had a large number of awards between 1990 through 2005. The selection resulted in: *Phalaenopsis* Brother Lawrence (25 descriptions), *Paphiopedilum micrantham* (92 descriptions) and *Sophrolaeliocattleya* Tiny Titan (44 descriptions).

The results of the tally suggest the difficulty of affirming my hypothesis (Table 1, page 218). The primary problem is an inconsistency in vocabulary to describe what are probably the same characteristic. For example, 24 of the descriptions of *Phal.* Brother Lawrence describe its substance as excellent, firm, good, hard, or heavy. The judges who wrote these descriptions probably saw (or felt) the same attribute, they just described it differently. However, judges apparently saw a different texture, with almost half (n=12) describing it as crystalline, diamond dust, or glistening; eight finding it matte or waxy, and one describing it simply as “good.” Similarly, most of the descriptions (n=21) of *Slc.* Tiny Titan suggest that its substance is average, good, medium or moderate. However, six judges found it hard or heavy, and 10 found it excellent or firm, suggesting some differences among cultivars. Descriptions of texture are even more variable, with 18 finding it crystalline, diamond dust, glistening, or sparkling; nine recording a matte

texture, and five finding a velvety texture. The consensus among judges is that *Paph. micrantham* has a substance that is average, good, medium or moderate (n=24), but others have found it crisp, firm, heavy, or light. More than half of the 32 recorded descriptions found a texture that is matte (n=17), but various others recorded textures that were average, crystalline, glabrous (smooth), hirsute, satiny, sparkling, velvety or waxy.

In summary, confirmation or refutation of the hypothesis is impossible because of the variability in vocabulary used by judges to describe substance and texture. The preponderant number of descriptions suggest that substance may be consistent across awarded cultivars of the same species or hybrid, but that texture may vary. However, because of variability of the descriptors used by judges, it is impossible to locate the variation either in the phenomenon being measured (the flower) or the measurement instrument (the verbal label).

This exercise convinces me that the description process would be well served if it included a checklist of descriptors — with definitions — to be applied to substance and texture statements. Don't the terms "crystalline" and "diamond dust" refer to the same phenomenon? Wouldn't our descriptions be improved if all judges used the same verbal label for that phenomenon? Table 1 suggests that there are probably five terms each for substance and texture that, if well defined, would cover the various differences found in flowers. Why don't we provide these to judges? More than half of the descriptions of *Paph. micrantham* omitted substance and texture statements altogether; perhaps these would have been included if a checklist with well-defined labels had been provided to description writers.

The structure of the AOS awards program suggests that the system benefits from reminders to judges about fundamentals in descriptions writing. Many of us found very useful the review in a recent *AQ* on the proper technique for taking a horizontal and vertical natural spread, especially on flowers with complex shapes.² More of these updates would probably find a welcome audience. The newly-created members' section of the AOS Web site would seem to provide an ideal venue for these updates and reviews.

·Hill, Cynthia, and Helmut Rohrl. 2001. Improving color identification of awarded orchids. *Awards Quarterly* 32(2):119–120.

·Roberts, Kenneth A. 2005. The proper method, taking horizontal and vertical measurements of awarded orchid flowers. *Awards Quarterly* 36(1):85.

Table 1. Substance and texture terms used in AOS award descriptions. (Based on descriptions published in *Awards Quarterly* of plants awarded from January 1, 1990 through December 2000.)

Descriptor	<i>Phal. Brother Lawrence</i>		<i>Paph. micrantham</i>		<i>Slc. Tiny Titan</i>	
	Substance	Texture	Substance	Texture	Substance	Texture
None	3	4	58	59	3	3
Average ¹			13	1	7	2
Crisp			1			7
Crystalline ²		7		2		5

Diamond-dust ³		3				3
Excellent	1		1		3	
Firm	5		3		7	
Glabrous				1		
Glistening						4
Good	1	1	7		2	
Glossy				1		
Hard ⁴	4				1	
Heavy ⁵	1		3		5	
Hirsute				2		
Light			1			
Matte		5		17		9
Medium ⁶			2		1	
Moderate ⁷			2		1	
Thin			1		4	
Satiny				4		
Smooth				1		
Sparkling		2		1		6
A slight sheen				1		
Velvety ⁸				1		5
Waxy		3		1		

KEY

¹Average. Includes "medium," "normal."

²Crystalline. Includes "lightly crystalline," "moderately crystalline."

³Diamond dust. Includes "jewellike." ⁴Hard. Includes "very hard."

⁵Heavy. Includes "exceptionally thick," "very heavy."

⁶Medium. Includes "medium to heavy."

⁷Moderate. Includes "fair," "above average," "typical for species."

⁸Velvety. Includes "glistening velvet."