doi: 10.1255/tosf.75

# The decision unit—a lot with objectives

Charles A. Ramsey
**EnviroStat, Inc., PO Box 657, Windsor, CO 80550, USA** chuck@envirostat.org

Sampling is more than shoveling material into a bucket. It is even more than using adequate mass, increments, and tools. Sampling is a systematic process that incorporates everything from development of objectives through final decision-making. Many sampling protocols currently in use focus only on the physical sample collection and ignore the preceding steps in the sampling process. The ignored steps include development of the critical decision objectives, integration of sufficient quality control, inferences from test portions to lots, and final decision making, statistical or otherwise. Without this supporting framework, it is impossible to ascertain the validity of the sampling protocols when needs or objectives change. Often, the same sampling protocol is implemented year after year without any consideration to its appropriateness.Proper Sample Quality Criteria (or Data Quality Objectives) are determined from the objectives of the project and must be an integral part of any sampling campaign. The major components of the Sample Quality Criteria are: 1) Question, 2) Decision Unit, and 3) Confidence. The Decision Unit is the specific material to which an inference from the analytical result is made and ultimately to which a decision is made. If the Decision Unit is not precisely determined and integrated into the development of the sampling protocol, the resulting decisions will be incorrect or, at a minimum, will not be cost effective. This contribution addresses development and integration of the Decision Unit into the sampling protocol framework.

## Introduction

The physical process of sample collection is a very complex endeavor. It entails the consideration of appropriate mass, number of increments, correct tools, randomness, maintaining sample integrity, etc. However, physical sampling is only a part of the entire process of making decisions with analytical data regarding a specific unit of material. The complete process includes developing objectives, understanding the nature of the material sampled, developing the sampling protocol, physical sampling (including sample processing and subsampling in the laboratory), interpreting the data, and final decision-making regarding the material in question, Figure 1.

The steps in the process (Figure 1) are briefly described below:

### Sample quality criteria

There are three parts of the Sample Quality Criteria (SQC)[1]

- Determination of the analyte(s) of interest and analyte concentration of concern. This is required to maintain analyte integrity and ensure proper care is taken during the sampling process not to contaminate the sample.
- Determination of the Decision Unit(s)[2]—the scale of decision-making. This will be addressed below.
- Determination of the confidence that the final decision is correct. This is a function of the error from the sampling process, how the analytical data will be used to make inference, and the consequences of an incorrect decision.

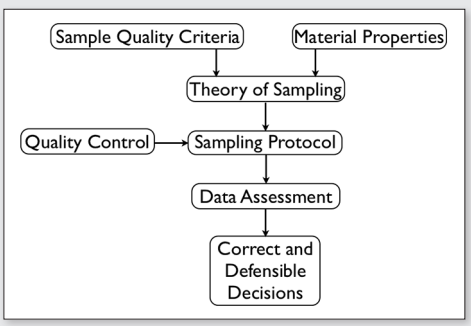


Figure 1. The comprehensive scientific, systematic process for defensible decision-making.

### Material properties

There are two primary material properties

- The nature of the elements. The elements may be finite (common with attribute type sampling schemes) or infinite (sometimes referred to as bulk materials). The Theory of Sampling (TOS) covers both types of elements though most effort is on the infinite element materials.
- The nature of the heterogeneity. This includes both the constitutional (compositional) heterogeneity and the distributional heterogeneity (in time and space).

### Theory of sampling (TOS)

The scientific principles that must be followed to develop a sampling protocol to ensure the samples meet the SQC.

### Quality control

The specific samples collected for the determination of error. Replicate samples are generally collected to measure precision (reproducibility). A variety of other quality control samples are collected (e.g., blanks to measure contamination) to ensure the sampling process is not introducing error.

### Sampling protocol

The specific instructions that must be followed to collect a representative sample (i.e., a sample that meets the SQC). It would address, among other items, sample mass, number of increments, selection and use of sampling tools, randomness, quality control, sample containers, necessary sample preservation, holding times, etc.

### Data assessment

The process of analyzing the data to determine if the criteria in the SQC are met—if the data is useful for decision making. A major component of data assessment is the estimation of the actual sampling error (from the quality control samples) and comparison of this

error to the error that can be tolerated. This is inversely proportional to the confidence desired.

### Correct and defensible decision

Once the data meets the SQC, they can be used to make *inference* (estimate the true concentration) to the Decision Unit. Once the true concentration is estimated, decisions can be made regarding the Decision Unit(s). These decisions could be to accept the Decision Unit as within specification, dispose of the Decision Unit (e.g. because it is contaminated), evaluate the Decision Unit further, etc. The list of potential decisions is almost infinite.

The objectives are as critical to the decision-making process as the physical sampling itself. Establishment of objectives is often overlooked. While Pitard addresses objectives[3,4] they are seldom developed by practitioners (at least in a manner suitable for the process of sampling). The objective that is most often overlooked and least understood is the scale of decision-making. The scale of decision-making, or the scale of observation, determines the specific material that needs to be included in the sample and the specific material the analytical results apply to. This scale of decision-making or observation is termed the Decision Unit (DU).

However, not all sampling objectives are related to making decisions regarding specific Decision Units. For instance, some sampling is performed for process control. However, in many fields including environmental, food, feed, pharmaceutical, chemical etc., testing products to determine if they meet specification or regulatory limits is very common. For this type of testing the Decision Unit must be established prior to sampling to determine if a limit is met. The Theory of Sampling uses the term "lot" to identify the material being sampled. Some common TOS definitions of the term "lot" are:

- The object to be evaluated[5]
- Batch of material from which increments and samples are selected[3]
- Sampling target, the specified material subjected to the sampling[6]
- All the material of interest[7]

Terminology is the cause of many disagreements and much frustration. It is therefore critical that terminology be very specific and precise so there is no room for misinterpretation. In the Theory of Sampling, the term "lot" is used to describe the material under investigation. However, this term (and the term population[8]) may not always be precise enough for development of sampling protocols and effective decision-making. In some cases, lot (and population) describes all the material under investigation, not the smaller amount of material that the decision is actually based on.

While these terms (lot, population, Decision Unit) appear very similar and descriptive, the following examples are given to demonstrate the limitation of how the term "lot" can be misused when making decisions.[†]

## Dog food example

A small pet food manufacturer is making dog food by mixing ingredients in a vessel (batch) that can hold 2,000 kilograms (kg) of dog food. The pet food is formed into kibbles. The manufacturer makes five such batches each day they produce this type of dog food. This type of dog food is manufactured approximately 20 days each year; therefore, approximately 100 batches of dog food are produced annually. The batches of dog food are placed in 10 kg bags for sale to retail customers. Depending on the size of the dog (and how many treats she gets!), a single bag of dog food may last one month.

In this scenario what is the lot? Is it the 100 batches, individual 2,000 kilogram batches, individual 10 kilogram bags, individual serving size or something else? The reader may already have an idea of what the lot is or may state: "that depends." If so, what does it depend on? The lot cannot be determined until the reasons for sampling (objectives) are developed. Incorrectly identifying the lot, or not indentifying a lot prior to sampling at all, are two very common sampling mistakes that must be discontinued!

The question that begins the Decision Unit discussion is the reason for testing the dog food. One reason may be to determine if the actual nutritional value of the dog food is the same as the nutritional value listed on the bag of dog food. In this case, the Decision Unit would be a bag of dog food. Another reason may be to determine if each batch has the same concentration (within a specified error) of some specific ingredient. In this case, the Decision Unit would be individual batches of dog food. Yet another reason may be exposure to potential toxins in the dog food. If one serving of dog food contains toxins above a certain level, the dog may develop a health issue. In this case the Decision Unit is a serving of dog food. In all cases it is the same dog food, but the scale of decision-making is different and therefore the sampling protocol would be different. What can complicate this even more is that not all analytes have the same Decision Unit. It could be for some analytes that an average over a 10 kilogram bag is compared to a nutritional limit and for a prohibited toxin, every piece of dog food (kibble) must be safe. The lot (or population) may thus be quite different than the Decision Unit.

The discussion of why the dog food needs to be tested determines what analytes to analyze the dog food for and what levels may be of concern, the Decision Unit, how the data will be used in the decision-making process, i.a. It is imperative that these discussions take place before the sampling protocol is developed, not after (which is common).

While the material in question may be a single 2,000 kilogram batch (should the batch be accepted or rejected), it may not be the average of the entire batch that is of concern but the percent of bags from the batch that have a specific characteristic. For instance, the batch may be deemed acceptable if 95% of the individual bags are within a certain specification limit. In this case, the Decision Unit is the individual 10 kg bags because decisions are made on the individual bags. This obviously has a large impact on the sampling

[†]The author is not advocating a change in terminology but rather an awareness of the use of the term in compliance (regulatory, specification, etc.) sampling. If the term lot is used with the same meaning as the term Decision Unit there is no conflict. In some industries, however, the term lot is used to define a specific amount of material with similar characteristics produced under like conditions. The lot number is very important for identification and trace-back of unacceptable goods in the food and feed realms for example. When working in such industries, it is very important to have a term different than lot for the material being sampled as the term lot has already been defined (many times in actual regulations). If not, confusion will result. The term this author and others have adopted for clarity and distinction from other terms is Decision Unit.[1,2,9]

Table 1. OSHA permissible noise levels.

| Duration per day, hours | Sound level dBA slow responses |
|---|---|
| 8 | 90 |
| 6 | 92 |
| 4 | 95 |
| 3 | 97 |
| 2 | 100 |
| 1.5 | 102 |
| 1 | 105 |
| 0.5 | 110 |
| 0.25 or less | 115 |

protocol. Any sample collected must represent an individual bag. If a sample is collected that represents the entire batch, it would be impossible to determine if the batch is acceptable because the percent of individual bags that meet the specification limit cannot be determined. In this case the 2,000 kg batch (material in question) may be viewed by some as a lot or population, but it is not the proper Decision Unit.

For a toxin, it may be a serving size (or daily amount) of dog food that is of concern. In other words, if a dog eats a serving of dog food that contains a toxin above a certain concentration, the dog may suffer some undesirable effect. In this case each and every serving of dog food must have a concentration of the toxin of concern below a certain level. In many cases this level would be the analytical detection limit. The Decision Unit is therefore each serving and there are many servings in each 10 kg bag. A sample that represents the entire 2,000 kg batch or even the 10 kg bag would not be sufficient to make a decision regarding the serving size.

## Noise level example

The US Occupational and Safety Health Administration has developed noise guidelines for worker exposures[10]. These guidelines state permissible average noise levels for specific length of exposure. For noise guidelines, the time of exposure is the Decision Unit. As with most exposure scenarios, the longer the exposure the lower the amount to which a receptor can be exposed. The eight-hour limit is 90 dBA, but the 30-minute limit is 110 dBA (Table 1). There are different Decision Units with different limits for each. If a reading is 100 dBA, is there a problem? There is no way to know unless the Decision Unit (time in this example) is specified as part of the measurement. Without a specification of the Decision Unit, it is impossible to interpret the data. If the value of 100 dBA represents a 30-minute Decision Unit, there is not a problem. If the value of 100 dBA represents a 4-hour Decision Unit, there is a problem. Would it be possible to determine worker safety unless information is known about all the Decision Units (nine of them) that exist?

## Coffee bean example

Coffee beans for import to the United States are regulated for mold. The current process to determine acceptance (conformance to the mold requirement) is to take 300 individual beans at random from the "lot" of coffee beans (usually beans are shipped in large sacks or containers). These beans are visually inspected individually for mold. If more than 25% of an individual bean is covered in mold, the bean is counted as moldy. If 21 or more of the 300 beans are moldy, the "lot" is submitted to the laboratory for further analysis. Otherwise the "lot" of coffee beans is accepted[11]. In this case, the Decision Unit is the individual bean and there are millions of these Decision Units. The analytical data (in this case a visual observation) applies to the individual bean. If a certain number (percentage) of these individual bean Decision Units meets a criteria, then the entire "lot" can be accepted.

The distinction of Decision Units and lot in this case is critically important. What if we have exactly the same testing and compliance scenario above, but the analyte of interest is not mold (attribute) but is some toxic compound (concentration)? This compound cannot be analyzed visually in the field, and a sample of 300 beans (minimum) is sent to the laboratory for analysis. Suppose the concept of Decision Units is ignored (forgotten, or never determined), and the laboratory decides to grind the entire sample so that a small portion can be selected for analysis (following the principles of TOS). In this case the analytical result represents the average of all 300 beans. This data could not be used to determine the percent of beans that have a certain threshold concentration, correct TOS or not. The result will be that a decision cannot be made, or the data will be used to make a decision, but that decision will be wrong (perhaps the correct decision will be made by dumb luck, but it would not be *defensible*). Determination of compliance is impossible without the concept of Decision Units.

## Exposure to toxic analytes

Exposure to toxic chemicals must also specify a Decision Unit (or Exposure Unit). Sometimes an upper concentration limit is stated for toxic analytes, and this limit is incorrectly used to determine future health risks without consideration of Decision Units (a very common scenario). An example may be lead exposure. Limits for lead in soil for residential areas are typically in the hundred parts per million range. For this example, we will assume a limit of 100 mg/kg (part per million) as the limit to determine if the soil is "safe" (the specific language varies from agency to agency) for residential use. The obvious question should be "what Decision Unit does this 100 mg/kg apply to?" Is the Decision Unit every gram of soil in the residential area, or all the soil the receptor is exposed to? The scale of the Decision Unit has large consequences in the design of the sampling protocol and the interpretation of the analytical results. In order to determine the scale of the Decision Unit, the model that was used to develop the 100 mg/kg limit must be understood. Does lead exposure come from a single gram that is over the limit or from all the soil the receptor is exposed to on a daily or annual basis. In the case of long term (chronic) exposure, it would be all the soil the receptor is exposed to during that time. For the case of short term/one time (acute) exposure, the Decision Unit would be smaller.

The sampling protocol must consider the Decision Unit or erroneous decisions will be made regarding the exposure of lead. For example, what if the sampling protocol is to collect one or more discrete (grab, specimen) samples and then subsample 1.0 gram for metals analysis? Can this approach determine the daily or annual exposure of lead to the receptor? The answer is obviously a resounding NO. The correct sampling protocol would be to collect increments (following the principles of TOS) across the same exposure area (could be space or time or both) used to develop the 100 mg/kg limit. Collecting samples from Decision Units that are smaller

than or larger than the Decision Unit used to establish the 100 mg/kg standard would be inappropriate.

## Summary of lessons from examples

- Decision Unit must be specified prior to the development of a sampling protocol.
- The Decision Unit may be specified in the case of compliance determination (regulatory or specification limits), or it may have to be developed (or determined) as in the case of exposure.
- The Decision Unit is the scale of decision-making which may be different than all the material in question. The material in question may be comprised of only one Decision Unit or many Decision Units.

## Conclusion

Development of the Sample Quality Criteria is critical for effective decision-making. Of all the components of the SQC, the determination of the Decision Unit is the least understood, yet it has the largest impact on the design of sampling protocols. The Decision Unit determines the scale of sampling, the scale of inference, and how data will be used to make decisions. Without a specified Decision Unit (which may or may not be synonymous with how the term "lot" is used), it is impossible to develop a defensible sampling protocol or to correctly interpret analytical data. Without knowledge and proper application of Decision Units, many incorrect decisions will be made.

The concept of Decision Unit is critical for the development of proper sampling protocols used to determine compliance (e.g., specification limits, regulations) as has been illustrated in the examples above. There are other terms used to identify the material under investigation, including lot, population, target material, etc. Sometimes these terms are not specific enough to indentify the specific material that the decision must be made on. The term Decision Unit identifies the specific material the increments are collected from and the specific material the results and decisions apply to.

## References


1. Ramsey, Charles A., Wagner, Claas, *Sample Quality Criteria*, Journal of AOAC International, Vol. 98, No. 2, March/April, 265-268 (2015).
2. Ramsey, Charles A., Hewitt, Alan D., *A Methodology for Assessing Sample Representativeness*, Environmental Forensics, 6:71-75, 2005.
3. Pitard, Francis F., *Pierre Gy's Sampling Theory and Sampling Practice* 2nd ed., CRC Press, 1993.
4. Pitard, Francis F., *Pierre Gy's Theory of Sampling and C.O. Ingamells' Poisson Process Approach*, Doctoral Thesis, Aalborg University, Denmark, 2009.
5. Gy, Pierre, *Sampling for Analytical Purposes*, Wiley, 1998.
6. *DS 3077 (2013) Representative Sampling– Horizontal Standard*, Danish Standards. www.ds.dk
7. Smith, Patricia L., *A Primer for Sampling Solids Liquids, and Gases,* ASA-SIAM, 2001.
8. Walpole, Ronald E., and Myers, Raymond H., *Probability and Statistics for Engineers and Scientists* 3rd ed., 1985.
9. Hawaii Department of Health, Office of Hazard Evaluation and Emergency Response, *Technical Guidance Manual for the Implementation of the Hawai'i State Contingency Plan*, 2009. http://www.hawaiidoh.org/tgm.aspx
10. U.S. Code Federal Regulations 29CFR1910.95(b)(2).
11. U.S., Food and Drug Administration, *Investigations Operations Manual*. http://www.fda.gov/ICECI/Inspections/IOM/default.htm