Sampling—is not gambling! (exit grab sampling)

Kim H. Esbensen[a] and Claas Wagner[b]
[a]Geological Survey of Denmark and Greenland (GEUS) and Aalborg University, Denmark. E-mail: ke@geus.dk
[b]Sampling Consultant. E-mail: cw@wagnerconsultants.com

The catchy title statement is attributed to the founder of the Theory of Sampling (TOS), Pierre Gy. It is a timeless response to the way most practical sampling is still being conducted today. IT is an elegant pun, but the statement of course needs substantiation. We here present examples of simple, easy-to-understand examples of *gambling*... instead of proper *sampling*. It is all about grab sampling and how this approach must be rejected with extreme prejudice. What follows is partly based on excerpts from Chapter 13: "Sampling–Hall of Shame", in a future textbook by Esbensen & Minkkinen, *Representative Sampling–in Science, Technology and Industry*.

Introduction

Panning for gold—there cannot be many better examples of an attitude of extreme *hope*. Perhaps the archetypal gold digger from the time of the great Gold Rush in the Western USA (latter part of the 19th century) illustrates the attitude indicated in Pierre Gy's statement in an optimal way. Unless you were both a gold digger and a qualified geologist with significant insight into the origin of placer deposits,† and most gold diggers certainly were not, the practice of gold panning was very much like gambling. Most gold diggers were trying out their luck in a specific area, along a specific creek… mainly because it had been rumoured that this was a lucky spot etc. Even along the same creek, the likelihood of finding gold nuggets is related to a quite restricted part of the full length only—heavy mineral nuggets travel down a river by saltation along the bottom and find their final resting place precisely where the hydrodynamic force of the water flowing along the bottom layer is no longer able to move the particles. Even with a geologist's professional knowledge and experience, panning for gold is still a somewhat risky endeavour. At the time of the gold rush, panning for gold was a very, very low probability gamble, but always with the greatest potential for a winning gambit lurking just out of sight. However, sometimes the gambit did indeed pay off: gold nuggets! In fact, one can pan for many other heavy minerals as well which, when concentrated enough, may also achieve favourable results.†

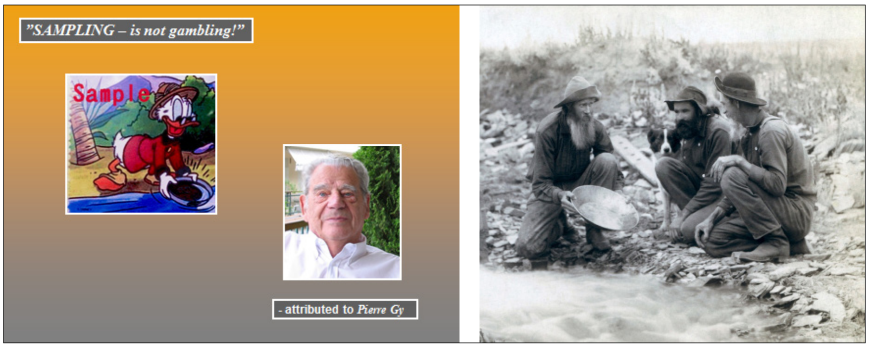

Figure 1. One of the most often quoted statements attributable to the founder of the Theory of Sampling, Pierre Gy. Gold panning is a wonderfully clear demonstration of "sampling based on hope", indeed a very clear form of gambling.

But for every winner there are innumerable losers… and so it is with sampling based much more on hope rather than on solid knowledge. Sampling—is not gambling!

†**Placer deposit:** mass movement and natural concentration of heavy minerals caused by the effect of gravity on moving particles. When heavy, stable minerals are freed from their matrix by weathering processes, they are slowly washed downslope into streams that quickly winnow the lighter matrix. Thus the heavy minerals become concentrated in stream, beach and lag (residual) gravels and constitute workable ore deposits. Minerals that form placer deposits have high specific gravity, are chemically resistant to weathering and are durable; such minerals include gold, platinum, cassiterite, magnetite, chromite, ilmenite, rutile, native copper, zircon, monazite and various gemstones.

Enough analogy

Enter the lessons learned in the first four sampling columns in *Spectroscopy Europe*: basic concepts and terms from the Theory of Sampling (TOS) necessary to appreciate the crucial role of lot/material *heterogeneity* and a first understanding that sampling processes *interact* with heterogeneous materials (and processes), which necessitate the understanding that sampling is first and foremost directed at *counteracting* the adverse effects of heterogeneity. The present column is all about how for every correct, representative sampling performed there are very many ill-reflected attempts at gambling,

Figure 2. Two manifestations of simple grab sampling—as applied to significantly heterogeneous materials or processes.

Figure 3. Compositional heterogeneity (CH) and distributional heterogeneity (DH) in real-world materials. How can one haphazardly selected "sample" (specimen) ever be assumed to be representative of an entire heterogeneous lot?

but with even, or lower, odds than the gold panner!

It is all about how grab sampling does not qualify in the light of heterogeneity. Grab sampling has already been introduced and commented upon heavily in previous columns, allowing us here simply to enjoy a series of examples that all have the hallmark of gambling—miles away from proper sampling (Figure 2).

There is one cardinal reason why grab sampling is the worst type of gambling: **heterogeneity**. Process streams, lots and materials in general that are significantly heterogeneous are so both with respect to compositional (CH) and distributional heterogeneity (DH), but the latter is the main enemy in all primary sampling stages (Figure 3). A singular, randomly selected extraction of material—meaning a grab sample—from a significantly heterogeneous lot cannot in any way, shape, weight or form hope to catch the characteristics of the entire lot, precisely because of DH_{lot}. Figure 4 shows that even three "back-to-back" samples can be very different indeed.

This principal understanding, laid out in schematic form in Figure 4, is illustrated with examples from the real world of materials in Figure 5. There is a very clear difficulty involved in any sampling procedure based on a single-extract operation (grab sampling).

"Well, take a bigger sample, then" is the suggested "remedy" most often heard. It is a completely wrong way of thinking, but we are obliged to take it seriously, in the didactic context of these columns. In Figure 6 is shown a very wide range of ever larger sample mass options (sample size), but it is clear that even a wheelbarrow will not necessarily catch a representative sample—it all depends on the magnitude of DH_{Lot}. Even IF a minor advantage *could* be obtained in the primary sampling stage with such an approach, there is only agony and despair waiting along the subsequent stages in the full "lot-to-analysis" pathway, where extraordinarily large masses are now required to be processed. We are in effect simply "passing the buck" but with no possible heterogeneity-counteracting solution at all.

Pierre Gy's famous statement is a reflection of the irresponsibility involved in hoping to obtain a representative sample without being willing to invest

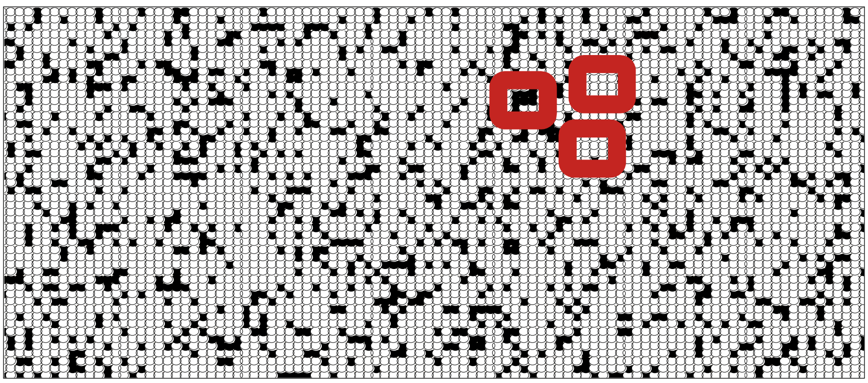

Figure 4. Any sampling process interacting with a material or a process stream that has the characteristics of significant distributional heterogeneity, DH_{lot} runs a fundamental risk of selecting and extracting worthless, non-representative **specimens**, a far cry from **representative samples**. The degree of risk is related both to the compositional as wel as the distributional heterogeneity in relation to the sample mass employed.

Figure 5. In addition to CH/DH issues, there may be equally severe grain-size heterogeneity issues for many types of materials. The right panel shows the utter futility of even trying to "cover" this type of heterogeneity, especially if the chosen (or mandated) sample size is manifestly too small for the job. Often a larger sample mass is claimed to be able to solve specific problems, but there is a very narrow limit to any potential benefits from barely increasing the proscribed sample mass.

Figure 6. No approach of employing "larger samples" will ever be able to counteract heterogeneity effects properly—yet there is still a clamouring for "bigger samples" in many walks of science, technology and industry. This approach will never eliminate DH, however, unless samples approach the **size of the whole lot**, obviously a preposterous notion. The focus is on the wrong entity—the issue originates with the heterogeneity not with the voluntary choice of the size of the sampling implement.

the necessary effort in learning a minimum of the principles in TOS, and simply continuing a long-standing tradition of cutting corners in the name of practicality, logistics, work effort and/or economy etc. These are considerations all pointing to the desire for grab sampling—because this approach is truly practical, and can always be carried out with a minimum effort, and which will therefore always end up as the cheapest "sampling" approach. Indeed grab sampling has all these desirable characteristics—and only one counteracting feature, albeit a most serious one: it can never be representative! Grab sampling amounts to a breach of due diligence. More, in-depth discussion on the merits (there are none) and the futility (unlimited) on even contemplating grab-sampling can be perused in Reference 1.

The next column will illustrate the only alternative to this situation, enter *composite sampling*.

Reference

1. K.H. Esbensen, C. Paoletti and P. Minkkinen, "Representative sampling of large kernel lots—I. Theory of Sampling and variographic analysis", *Trends Anal. Chem. (TrAC)* **32,** 154–165 (2012). doi: http://dx.doi.org/10.1016/j.trac.2011.09.008

Kim H. Esbensen originally trained as a geologist/geochemist, but it was 30 years before he actually worked in a geoscience institution (The Geological Survey of Denmark and Greenland). In between he established two research groups dealing with PAT and chemometrics. He found a third love, scientifically speaking, some 15 years ago, when he met the Theory of Sampling (TOS), and the field of representative sampling has occupied his career ever since. Kim is specifically interested in the interaction between process and material heterogeneity, representative sampling and augmented measurement uncertainty.

Originally trained as an economist, **Claas Wagner** realised that his real interests were with environmental and energy related topics and therefore continued his education in this direction. Sustainable resource management, emission reduction procedures and energy efficiency issues have all one common ground: decisions need to be based on valid data. This led to Claas' PhD on representative sampling and data analysis for quality monitoring in large-scale combustion plants. Currently Claas combines his fields of interest, working as a consultant for various industries providing quality assurance approaches. Throughout all of this reigns representative sampling.