

QUALITATIVE SAMPLING v STATISTICAL ANALYSIS v DAUBERT

By Lonnie Haughton

An increasingly common analytical error is the assumption that the accuracy and integrity of an expert's evaluation of building envelope construction or performance can be judged via probability analysis. It may be argued that the only valid approach for forensic surveys is the blind "simple random sampling" (SRS) that many of us were taught in our high school or college science classes. Alternatively, it may be asserted that the expert should have employed an advanced "stratified random sampling" design in which various components or characteristics of the building envelope were identified in advance of the testing, followed by random sampling within these distinct "strata."

To better evaluate these claims, it is helpful to list key elements of the basic SRS protocol:

- Sampling locations are predetermined randomly. The sampler has no discretion in the selection process;
- Every member of the "population" being sampled has the same chance to be selected; and
- The results of any particular sample cannot be used to shape the course of the continued sampling.¹

Upon only brief review, experienced readers of *Interface* will identify potentially critical hindrances to random probabilistic sampling of buildings, including:

- Aesthetic or logistical constraints imposed by the building owner or occupants;
- Physical or legal limitations on access; and
- The inordinately high costs imposed by the inflexible sampling design.

Further, any SRS design requirement that data obtained during the course of the sampling cannot be used to guide the immediate additional sampling represents a flagrant waste of the building envelope professional's forensic expertise.

In a sense, such misjudgments are a testament to the general public's blind abiding faith in the inerrant power of statistical formulas and calculations; however, the proponents of statistical sampling also may simply lack a practical understanding of how building envelopes are assembled. In particular, any statistical model developed years after a building's construction that is claimed to be sufficiently robust to account for the generally nonrandom nature of this work instead may simply demonstrate the limited knowledge of the model's creators.

For example, consider the installation of fluid-applied deck waterproofing by a small

crew during a single eight-hour workday, and assume that on previous days, this team already has waterproofed other decks at the same project. It is safe to predict that on this particular day, the manner and quality of the waterproofing carried out at various decks by this crew will be generally consistent. Further, barring recently revised instructions from their superintendent or other inspectors, it is reasonable to assume that the team members' output on this particular day will be generally consistent with their output of the previous day.

Yes, there will be inconsistencies and deviations (for better or for worse) in the intertwined streams of work that produce a completed building. And yes, the composition of the project's crews may change at irregular intervals with unforeseen effects. During the extended process of constructing a building, however, a generally consistent day-to-day level of workmanship (whether good, middling, or deficient) is the norm. Years later, after the workers have dispersed, their managers' memories have faded, and the availability of project records is limited, any attempt to evaluate via probability analysis the patterns (or the blips in the patterns) of their outputs likely will be inaccurate, incomplete, and perhaps biased.

In contrast, the industry standard for forensic evaluations of building envelope

performance is ASTM E 2128, *Standard Guide for Evaluating Water Leakage of Building Walls*,² which prescribes a purposeful investigation that entails an orderly accumulation of information in such a manner that each step enhances and supplements the information gathered in the preceding step. Clearly, the sampling methodology described within ASTM E 2128 is not consistent with the tenets of probability sampling and its various statistical siblings (collectively, quantitative sampling), as exemplified by the SRS ideal, and therefore may be subject to claims that it does not satisfy the rules of evidence that require an expert's sampling methodology and testimony to be based upon "scientifically valid" principles (e.g., *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S., 579, 1993).

In response, this writer and the well-known consultant Colin Murphy, RRC, FRCI,³ have published a peer-reviewed paper⁴ in the *Journal of ASTM International (JAI)* to demonstrate that the survey protocol prescribed by ASTM E 2128 is fully consistent with qualitative sampling methodology validated within the social sciences. We note that Section 11 of ASTM E 2128

includes, in part, the following protocol for carrying out a survey of the building envelope:

1. An evaluation is conducted in response to a problem situation and a nonperforming wall and may involve several techniques and procedures specifically adapted and applied in a systematic manner to diagnose a specific problem.
2. The information systematically accumulated in a leakage evaluation is analyzed as it is acquired. The new information may motivate a change in approach or focus for subsequent steps in the evaluation process.
3. The evaluator is expected to establish a cause-and-effect relationship between wall characteristics and observed leakage. This requires an appropriate selection of activities and a logical analysis and interpretation of the acquired information.
4. The conclusions and findings from an evaluation must be rationally based on the activities and procedures undertaken and the information acquired if they are to be considered legitimate and substantiated.
5. The record should be sufficiently complete so that any interested party can duplicate the evaluation program and acquire similar information. Notes on the analysis and interpretation of the acquired information should be clear and complete enough to be understood by any other building professional skilled in leakage evaluation.

Clearly, the authors of ASTM E 2128 consider the proper role of the skilled forensic professional is to carry out a nonrandom (i.e., qualitative) investigation. To better evaluate this position, we turn to seminal standards⁵ in the social sciences that review the different logics that undergird the quantitative and qualitative sampling methodologies. While quantitative methods generally require a large, randomly selected sample set, qualitative inquiry typically focuses in depth on relatively small samples, even single cases (N=1) selected purposefully. Not only do the sampling techniques differ, but the very logic of each approach is unique because their purposes are different:

- The logic and power of random sampling derive from statistical proba-

bility theory – a random and statistically representative "sample set" controls selection bias and permits confident generalization from the sampling to a larger population.

- In contrast, what would be "bias" in statistical sampling, and therefore a weakness, becomes intended focus in qualitative sampling, and therefore a strength. The logic and power of purposeful sampling lie in selecting information-rich samples for step-by-step evaluation of issues of central importance to the purpose of the inquiry.

Consider the accompanying photographs of a three-story, 120-unit wood-framed apartment complex facing a central courtyard and constructed above an at-grade concrete parking garage with busy commercial spaces (pizza restaurant, coffee shop, etc.) located at the building perimeter. Assume that the building owner has reported water damage at various deck soffits and at interior window surrounds and that an initial visual survey indicates that at the residential levels, there are 750 window openings and 80 wood-framed decks with a concrete walking surface over fluid-applied waterproofing. Also assume the owner has requested that the leakage investigation be carried out in a manner that does not impact the operations of the commercial shops at the building perimeter (see *Photo 1*). Further, let's assume there are compelling reasons to focus the bulk of the destructive testing on the readily accessible units at the interior courtyard (see *Photo 2*).

In other words, let's assume that 65% of the residential windows and privacy decks at this apartment complex are unavailable for destructive testing. These limitations might be daunting to random samplers, but not for the experienced investigators who recognize the strong likelihood that there will be a generally consistent level of workmanship at all of the decks and windows. This critical assumption does, of course, need to be closely evaluated during the course of the ensuing investigation (including any further review of available construction records); however, it encourages the investigators to commence their destructive testing at any convenient location that can be expected to be information-rich.

Broadly generalized, there are three categories of sampling locations selected by most building envelope professionals:

1. Category 1 sampling locations typi-



RCI, Inc.
800-828-1902
www.rci-online.org



Photo 1 – Minimal testing occurred at this streetside elevation due to a mandate to avoid disturbing the operations of the coffee shop and a nearby pizza restaurant.

Photo 2 – Instead, the majority of the qualitative testing was carried out at the more readily accessible central courtyard.



cally occur at changes in material and changes in plane, such as fenestration and deck-to-wall transitions, that can be expected to be highly information-rich due to the number of waterproofing, flashing, sealing, and/or water-resistive barrier transitions by various trades that can be exposed in sample areas of relative limited size. Most expert investigations will commence at readily accessible, information-rich locations, often at weather-exposed elevations.

2. Category 2 sampling locations, such as vent penetrations and handrail connections, are transitions that likely are not as information-rich as the Category 1 locations, but which can provide key supplemental information regarding consistency of design details, quality control, potential extent/severity of any deterioration, and quality of the work carried out by some of the subtrades.
3. Category 3 locations are those that are sampled to address atypical construction or design issues or specific questions or issues of concern that may arise during the course of the sampling. Category 3 sampling often is carried out in the later phases of the survey process and can serve various survey closeout pur-



RCI, Inc.
800-828-1902
www.rci-online.org

It is important to recognize that qualitative and quantitative sampling methodologies constitute alternative, but not mutually exclusive, strategies for investigative research.

poses, including localized sampling intended to test alternative conclusions raised by others and purposeful random sampling to increase the perceived credibility of the qualitative analysis.

While the goal of quantitative sampling is to evaluate levels of statistical significance, the methodology of qualitative inquiry is to produce findings that have substantive significance, which refers to the strength and importance of a meaningful relationship.⁶ In determining substantive significance, both the analyst and the reviewers must address these kinds of questions:

- How solid, coherent, and consistent is the qualitative evidence in support of the expert's findings?
- To what extent and in what ways do the findings deepen understanding of the observed conditions?
- How well do the findings define and correlate causal relationships in a manner that maximizes understanding of the various processes and phenomena of interest that are occurring within the population?
- To what extent are the expert's findings consistent with knowledge derived from other sources?

The sampling design for qualitative surveys typically is fluid, capitalizing on early learning to guide subsequent direction. Qualitative researchers begin by identifying information-rich data sources that can be expected to maximize understanding of the observable conditions. As the survey progresses, new conditions and sampling ques-

tions emerge that may confirm, enrich, modify, or challenge the researcher's understanding of the observed phenomena. Qualitative sampling represents an inductive process in which the researcher searches for patterns and builds abstractions, concepts, hypotheses, and theories from emerging details.

Unlike probability sampling, there are no firm criterion for determining sample size in qualitative surveys. The appropriate sample size is determined by the quality of the observed data as they relate to the goals of the survey. If the key purpose is to maximize information, then the survey may be terminated when no new information is forthcoming from the additional sampling; thus, "data saturation" may be the primary criteria. However, sample size for qualitative surveys also may depend simply on what one wants to know, the purpose of the inquiry, what will be useful and will have credibility, and what can be done with available time and resources.

At the apartment complex seen in *Photos 1* and *2*, let us assume the plaintiff's destructive testing of the concrete walking surface at four of the leaky decks at the central courtyard has revealed voids (see *Photo 3*) in a nonreinforced, fluid-applied waterproofing membrane that measures

less than 1/16 in thick. Then, upon review of the product's application instructions, we learn that the manufacturer recommends a minimum 1/8-in thickness and an embedded layer of fiberglass reinforcing mesh. Let's further assume that dropping the soffits at eight other courtyard decks revealed water-damaged OSB decking consistent with similar breaches in the waterproofing membrane.

At this stage of the qualitative survey, there is sufficient substantive significance to this accumulated data to make a legal case and to allow the plaintiff's expert to make a preliminary assertion that all 80 decks are deficiently waterproofed. At this point, the ball is in the defense experts' court, where it is a simple matter to confirm or discount this extrapolation by opening concrete at two more decks and dropping the soffits at several others (including one deck at the building perimeter).

In summary, the entire investigation (by plaintiff and defense) of the deck waterproofing at the 80 wood-framed decks at this project consisted of limited concrete removal at six decks and opening the soffits at 11 or 12 others. All of the test locations were purposefully selected in accordance with industry standards and the customary practice of experienced professionals. Because the overall results of these two

qualitative investigations were consistent, any further sampling by either party would have been a waste of resources.

On the other hand, if the survey results had been inconsistent, further qualitative sampling may have been needed in order to achieve the good-faith goal of nonbiased consensus regarding the installation and performance of the fluid-applied deck waterproofing. In any case, at these decks there was no appropriate role for statistical sampling by the plaintiff or the defense; any later claim to the contrary was simply a tactical diversion by defense attorneys seeking to better their position.

Even so, it is important to recognize that qualitative and quantitative sampling methodologies constitute alternative, but



Photo 3 – Typical breach in the overly thin, nonreinforced, fluid-applied waterproofing membrane.

not mutually exclusive, strategies for investigative research. The experienced, pragmatic investigator practices a situational responsiveness that recognizes that differing methods and techniques are appropriate for different circumstances.

In the end, the core of the scientific method for all quantitative and qualitative surveys consists of a puzzle-solving strategy or method for analysis (or elimination) of rival explanations or hypotheses. This strategy may start its puzzle solving with a hypothesis (i.e., quantitative analysis) or it may start with evidence (i.e., qualitative analysis). The quantitative survey begins with the formation of a hypothesis that can be evaluated statistically upon later collection of evidence, while the qualitative survey begins with the collection of evidence from which substantive explanations will emerge. In both cases, the core of the scientific method is represented by the strategy of analysis of plausible rival hypotheses.⁷

As the final step in the qualitative survey process – after describing and interpreting the major patterns, themes, and linkages that have emerged from the analysis – the expert investigators (plaintiff and defense) must, as a matter of intellectual integrity, look for data that support alternative themes and explanations for the observed conditions. Failure to find strong supporting evidence for opposing theories increases confidence in their original analysis. Closely related to such testing of rival explanations is the search for negative cases: comprehensive understanding of the perceived qualitative patterns is increased by considering the instances and cases that do not fit the pattern.⁸

Permeating the building envelope survey strategy delineated in ASTM E 2128 is the critical goal of tying together cause(s) and effect(s). Within the limits of the investigator's commission:

1. The consequences of leakage are established;
2. The severity, consistency, and distribution of these consequences are determined;
3. The leakage pathways are defined; and
4. The sources of these building envelope failures are identified.

This investigative process commonly is both inductive and deductive and should be carried out with methodological competence, intellectual rigor, and professional integrity. The pragmatic investigator will


implement a range of qualitative and quantitative measures that best evidence credibility when reviewed by the target audience. For the well-trained building envelope professional, general conformity to the survey protocols published within ASTM E 2128 constitutes a form of analytical rigor comparable to the validated tenets of qualitative analysis practiced within the social science fields. If this survey process and its findings are reviewed during the course of litigation, the rules of evidence clearly have been satisfied when a building envelope expert's sampling methodology and analysis are

founded upon ASTM E 2128.

It is this writer's experience that an analysis of legal efforts to use probabilistic analysis to discredit the results of an expert's qualitative survey will reveal critical flaws in the assumptions that undergird the critic's statistical model. Opposing attorneys would be well advised to closely examine the statistician's knowledge of the day-to-day processes of constructing the building envelope. The analyst who claims the ability to craft a statistical model of a complex, nonrandom series of unfamiliar and poorly documented events that occurred



RCI, Inc.
800-828-1902
www.rci-online.org

years earlier is most likely simply naïve, but he or she may also lack the methodological competence, intellectual rigor, and professional integrity required of a legal expert in accordance with Daubert standards. 

Editor's Note: A shortened version of this article was published in the April 2008 quarterly newsletter of the Forensic Expert Witness Association (www.forensiv.org).

FOOTNOTES

1. D. Freedman, R. Pisani, R. Purves, and A. Adhikari, *Statistics, Second Edition*, W. W. Norton & Co, New York, 1991.
2. ASTM International, www.astm.org.
3. Colin Murphy, RRC, FRCI, is the founder and managing principal of Trinity | ERD (www.trinityerd.com), a building envelope forensics, testing, and design consulting firm based in Seattle, WA.
4. L. Haughton and C. Murphy, "Qualitative Sampling of the Building Envelope for Water Leakage," *Journal of ASTM International*, Volume 4,

Issue 9 (October 2007).

5. In particular, M.Q. Patton's *Qualitative Research & Evaluation Methods* by Sage Publications was an invaluable reference resource for the paper listed in reference 4 above and for the qualitative research methodology discussed within this article.
6. W.P. Vogt, *Dictionary of Statistics and Methodology*, Sage Publications, 1993.
7. Credit for this insightful perspective

belongs to M.Q. Patton and also to Donald Campbell from his acclaimed foreword to R. K. Yin's book: *Case Study Research: Design and Methods* by Sage Publications.

8. Note again that M.Q. Patton's *Qualitative Research & Evaluation Methods* by Sage Publications is an invaluable reference resource for the qualitative research methodology discussed within this article.

Lonnie Haughton

Lonnie Haughton, CDT, LEED AP, is a construction consultant with Richard Avelar & Associates, a forensic architectural consulting firm in Oakland, CA. His professional Web site is www.mastercodeprofessional.com. Lonnie is one of the fewer than 500 individuals nationwide who have achieved the Master Code Professional certification awarded by the International Code Council. He is a member of the Forensic Expert Witness Association, the Western Construction Consultants Association (Westcon), and the Construction Writers Association.

