Figure 3.5. SAMPLING: Limitations Imposed by Sample Error – Two Illustrations

EM9010: Kitchener-Waterloo Record, January 19, 1990, pages B1, B2

Schneider's won't get rebate over blood spilled into sewer

By Monica Gutschi
Record staff

It was "a coincidence" that Waterloo Region tested J.M. Schneider Inc's waste water discharge just after several accidental blood spills into the sewage system last spring, the company's executives said Wednesday.

"This period of time is highly unusual", said Bob Tiffin, the national quality assurance manager for the Kitchener meatpacking company. "It's not typical of our history and our operation."

But that didn't convince Waterloo Region's engineering committee, which refused to rebate any of the $222,970 sewage charge Schneider's paid for the discharged wastes.

The committee split 4-4 on a motion to give Schneider's a $73,354 rebate on its sewage charge, which meant the motion was defeated. Earlier it rejected a staff recommendation to give the company a $156,708 rebate.

Regional council will discuss the issue at its next meeting and may take a different stand.

"We don't hear about spills unless they occur when we have tested", said Kitchener Ald. Jim Ziegler. "There are other times that we don't know about that have affected the effluent. There has been since September 1987 numerous spills that haven't occurred near testing days."

Four tests conducted on the company's waste discharge in April, May and June last year showed average biochemical oxygen demand of 2,178 parts per million and an average suspended solids measure of 1,412 parts per million. The historical average for the company is an oxygen demand of 675 ppm and suspended solid measure of 550 ppm.

Schneider's regularly pays a surcharge when its discharges into the sanitary sewer system exceed the region's limit for suspended solids (TSS) and biochemical oxygen demand (BOD) – a measure of the oxygen the region needs to supply to reduce the waste in the effluent when it reaches the Kitchener sewage treatment plant.

Two of the tests were taken on the same day that a blood spill had occurred while the other two tests were taken the day after a spill. There were also two other spills in that time period.

"It appears to be a coincidence that the spills occurred on the regional testing days", Tiffin said. "The high-strength discharges were not continuous for the entire quarter."

Schneider's had asked for a rebate, arguing the surcharge assumes high levels of waste in the effluent for the entire quarter, rather than just on the testing days.

The request was supported by members of the engineering department staff.

On those (testing) days, the company had manufacturing problems which overwhelmed the system", Ralph Luhowy, waste-water operations manager, told the committee. He said the company would not keep operating while spilling blood because it is used in a number of products.

Tiffin said Schneider's has begun a daily sampling program of its sewage discharge and has purchased a blood-holding tank to avoid the overflows that caused the spills.

He said company executives were "surprised" to find out about the blood spills, the failure to report them and the failure to communicate with the regional laboratory.

Waterloo Coun. Susan Forwell said that while she was willing to give the company the "benefit of the doubt", she would not do so the next time it had problems.

1 Identify the *measure(s) of location* involved in the dispute; also give their *value(s)* and their *unit(s).*

2 What category of error in the reported values of the measures of location is the primary source of the uncertainty that is the basis of Schneider's claim that the $222,970 sewage surcharge was too high?
- What measure(s) of this error are given in the article?
  - Explain briefly what other summary statistics should be given in the article as well as the four values of the averages.
- What *other* category of error (not mentioned in the article) is an additional source of the uncertainty?

3 How is *inaccuracy* in measuring biochemical oxygen demand and suspended solids addressed in the article? Comment briefly.
- Explain briefly how the inaccuracy of these measurements could be quantified?

4 The article states that four tests were carried out over a 91-day period. Assuming that spills occurred equiprobably over the period and that there were only six spills in total, calculate the probability of results *as extreme or more extreme* (from Schneider's perspective) as those observed; *i.e.*, 2 tests falling *on* the day of a blood spill, 2 on the day immediately *after* [*more* extreme results are (3,1) and (4,0) for the two numbers]. Show your reasoning.
- Assuming a value of about 0.0001 for this probability, what do you infer about:
  - the days chosen for testing by Waterloo Region;
  - the statement that only six spills occurred over the 91-day period?

5 What would you recommend if you were a statistical consultant helping Waterloo Region avoid such a dispute in future?

6 What do you infer about the internal organization of J.M. Schneider Inc. from the second-last paragraph of the article: *He said company executives .... with the regional laboratory*? Explain briefly.

(*continued overleaf*)

Matters of statistical interest raised by the second article (reprinted below) in this Figure are as follows.

* The limitation, highlighted below in the middle of the page, imposed by sample error on an Answer to a Question of interest – in this instance, the sample size of *one* exacerbates the severity of the limitation.
* The use of equiprobable (or random) assigning (EPA) of the 200 toys to the 20 groups of 10 and the use of equiprobable selecting (EPS) for the two toys the dog, Rico, was asked to fetch in succession from each group of 10.
  – A response variate in this investigation was the number of words Rico knew; one Answer is presented in terms of a *binary* response variate – each of Rico's choices was *right* or *wrong* – and the attribute of interest was then the *proportion* of right choices (*e.g.*, 37 out of 40) in an equiprobably-selected subset of (40 of) the 200 possibilities.

EM0420: The Globe and Mail, June 11, 2004, pages A1 and A14

Dog impresses scientists with 200-word vocabulary

Scientist noticed dog on popular game show

BY ANNE MCILROY

He may be a dog genius. But a Border collie named Rico with a vocabulary of 200 words is challenging the belief that humans have a unique way of learning language.

German scientists, who discovered Rico on a game show, say he quickly learns new words, such as sea horse, tiger or panda, the same way toddlers do.

He appears to pick up new words after hearing them only once, a process called fast-mapping, which allows children to add words to their vocabulary even if they just overhear an adult using them. Up until now, researchers had thought this was a uniquely human ability.

No, Rico doesn't speak.

"You don't have to be able to talk to understand a lot," says Julia Fischer, a German researcher with the Max Planck Institute for Evolutionary Anthropology in Leipzig.

Her research appears to back up dog owners, especially those with Border collies, who say their pets have a knack for learning words.

"I picked a Border collie because I wanted one little creature in my life to listen to me," says Evelyn Cheesbrough, an Ottawa dog owner with two children.

Dr. Fischer and her colleagues caution that their work involves a single dog. It could be that nine-year-old Rico is brilliant, or that he belongs to a breed with an exceptional ability to learn words, or that the early training he received from a German family makes him different from other dogs.

Since he was 10 months old, his owners have trained him to fetch more than 200 items, mostly children's toys, upon request.

Dr. Fisher noticed Rico on a popular German game show in which he surprised viewers with his large vocabulary. She approached his owners and began her research by verifying that the dog did indeed understand 200 words.

She put the 200 items with which he was said to be familiar into 20 random sets of 10 toys each. For each set, Rico was asked to fetch two randomly-chosen toys – for example, a dinosaur and a red ball – from another room from which he couldn't see his owner.

He got 37 out of 40 right. This gave him a vocabulary comparable to that of apes that have been trained for years to understand language, and to dolphins and parrots.

But the next two experiments put him on par with humans, at least very young ones. Rico was asked to fetch an unfamiliar item

It could be that nine-year-old Rico is brilliant, or that he belongs to a breed with an exceptional ability to learn words, or that the early training he received from a German family makes him different from other dogs.

– for example, a bunny – that was added to a group of seven familiar toys. He was able to pick out the new toy seven times out of 10, apparently because he understood that the familiar items already had names.

Four weeks later, he was tested to see if he remembered the new objects. Could he pick out the bunny, choosing from a set of toys that included four completely new toys and four familiar ones? In six tests involving different new toys, he picked the correct item half the time, comparable to the performance of three-year-olds in similar experiments. Tested 10 minutes later, he got it right four times out of six.

The findings, which were published in today's edition of the U.S. journal *Science*, suggest that brain structures that support this kind of learning may not be unique to humans.

There are, however, clear differences between language acquisition in humans and dogs. Toddlers have a much broader knowledge than Rico about the meaning of words, the researchers say, and can distinguish between words that convey meaning on their own, such as nouns and verbs, and those that don't, such as articles.

Rico also can't speak, and at nine years old knows just 200 words. Human nine-year-olds know tens of thousands of words, and are learning 10 new ones a day, says Paul Bloom, a psychologist at Yale University who wrote a companion piece on Rico for *Science*.

More experiments are needed. Dr. Fischer and her colleagues are investigating whether Rico can understand entire phrases, for example, "Put the bunny in the box." In an interview, she said her team is planning to study the vocabulary of other Border collies and other breeds of dog.

Not all may have the same linguistic aptitude. An unscientific survey of dog owners at an Ottawa park found that their pets know anywhere from 37 words to none at all.

Warren Coutts owns two greyhounds, former racing dogs he adopted. He says they understand just a few words, but are smart, and quickly adapted to a quiet home life as pets. "They are smart in a socially adaptive way."

Jeannie Wynne-Edwards isn't sure her Great Dane, Daisy, understands any words. She used to think that was because Daisy grew up in a francophone family before moving to her. She has tried speaking to her in French, to little effect.

"She's ineffective in both languages."

REFERENCE: Kaminski, Juliane, Call, Josep and Julia Fischer: Word Learning in a Domestic Dog: Evidence for "Fast Mapping." *Science* **304**(#5677): 1,682-1,683, June 11 (2004). [UW Library E-journal]

See also Bloom, Paul. Can a Dog Learn a Word? *Science* **304**(#5677): 1,605-1,606, June 11 (2004).