

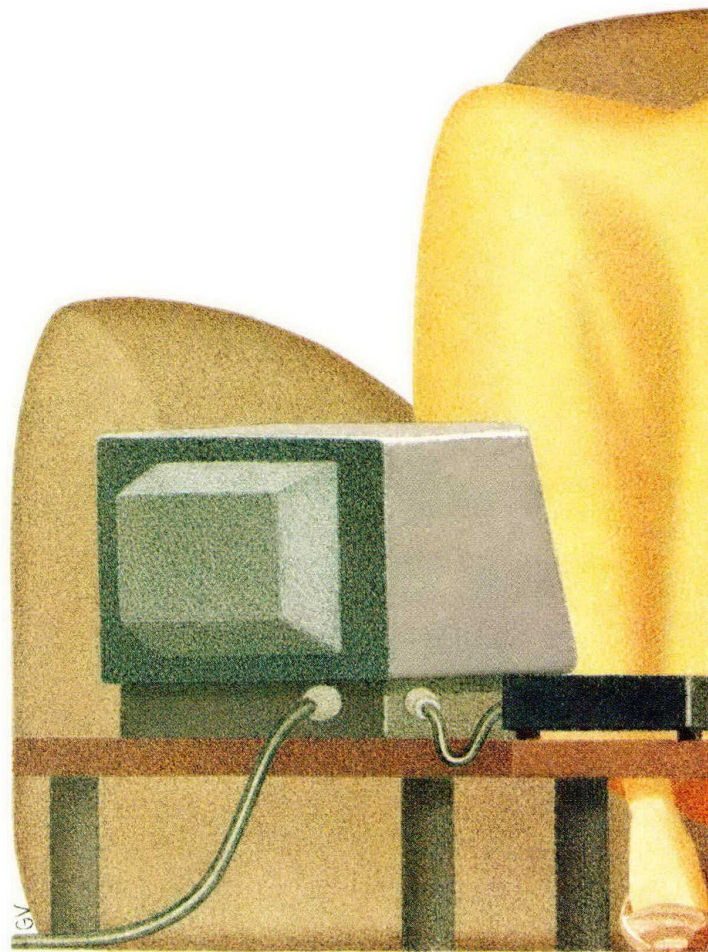
A remarkable new technology may soon tell us precisely who watches television, what they watch, and for how long—and thus require a major restructuring of the way television conducts its business

WATCHING AMERICANS WATCH TV

BY ERIK LARSON

A GREAT MYSTERY HAS GNAWED AT THE HEARTS OF advertisers ever since the Bulova Watch Company paid for the first TV ad, which aired on July 1, 1941: How many people stick around for the commercials? Advertisers have always had to come at the question from an oblique angle, measuring the audiences for the programs surrounding their commercials rather than for the commercials themselves. The A. C. Nielsen Company brought technology to the business of estimating audience size in 1942, with the first mechanical meter, the Audimeter, which used a stylus to scratch out a record of where a radio dial was tuned. The system was crude but once refined proved far faster than the competition's surveys, and Nielsen quickly gained monopoly control over the business of national ratings, at first for radio and later for television. Most viewers know Nielsen as the maker of the bullets that killed *Star Trek* and *Twin Peaks*, but to think of its ratings exclusively in terms of their show-stopping power is to underestimate the depth of Nielsen's influence over the culture, content, and business of television, and thus over the evolution of our consumer culture itself. Nielsen *is* television. Imagine a company like IBM granting to an outside party a decisive say in whether it has a profit or a loss in a given quarter, and you'll begin to understand Nielsen's peculiar hold over broadcasting.

Nielsen kept competitors at bay through the deft and timely use of technology, but each fundamental advance changed the landscape of television. All along, the com-



pany's engineers pursued TV's Golden Fleece, the elusive "passive" audience meter, a device capable of recognizing each TV viewer and recording even his briefest exit, without his having to lift a finger. Audience engineers have dreamed of such a device ever since that epiphanic instant in 1952 when a Toledo, Ohio, water commissioner realized that a sudden city-wide drop in pressure coincided exactly with a commercial break in *I Love Lucy*, and joked that someone should invent a "flushometer." The men from Nielsen tried weight-sensing scales and electric eyes and even sophisticated military technology of the kind we so vividly saw put to work during the Gulf War. Still, some insurmountable obstacle always seemed to get in the way—big dogs, for example, the bane of any audience engineer.

Until lately Nielsen could afford to pursue this quest at



a leisurely pace. Now, however, Nielsen's three largest customers, ABC, CBS, and NBC, openly charge that the company's existing "people meter" technology is unreliable, perhaps obsolete. After an exhaustive investigation they concluded in 1989 that their \$10 billion industry had become a helpless captive of some rather flimsy statistics. Soon Nielsen reported steep, inexplicable declines in key segments of TV viewing which seemed to confirm the networks' worst fears and deeply shook their faith in the company's ratings.

"The thing we've always had has been this set of numbers," says Alan Wurtzel, ABC's senior research executive. "It's the foundation upon which program decisions are made, deals are made, advertising dollars are spent. When those numbers became unreliable, inconsistent, inaccurate, unexplainable, they just threw into chaos almost

everything we did, because they were the basis on which the business was built."

To make matters worse, last April the networks invited all comers to present ideas for alternative ways of measuring the U.S. audience, the television equivalent of having your spouse suddenly begin advertising for a new lover. The competition has begun to stir. The Arbitron Company, which competes fiercely with Nielsen to rate local TV markets, continues to threaten to go national; it has funded an effort at the Massachusetts Institute of Technology's Media Lab to develop a passive meter and last November began tracking public response to the full network schedule. A French company has already deployed a kind of passive meter; Arbitron has begun testing the system for use here. Britain's AGB International, which was deeply gored in a past encounter TVision Insights, Inc. — Ex. 1016, Page 000002

with Nielsen, hints that the U.S. market is simply too lucrative to resist.

Nielsen, therefore, finds itself virtually forced to deploy its newest secret weapon—a device that may be the ultimate passive meter. I met the machine in New York. It can identify viewers and even gives a rough indication of whether their heads are turned toward the TV. The system could fix a lot of the problems cited by the networks, but its deployment would have a significance far beyond its capacity to produce accurate numbers. By tracking the TV audience on a second-by-second basis, Nielsen's passive meter may yield the first accurate ratings of commercials—not shows, *commercials*—and thus at last solve that great nagging mystery about the “reach” of advertising. If it does, however, it will sharply alter the way consumer products are marketed in America, even the way television shows are written and produced. “The implications,” says John Dimling, Nielsen Media's executive vice-president, “could be quite staggering.”

But does anyone really want to know the truths that Nielsen's black box (and for now it really is a black box) may bring? What's wrong with Nielsen's numbers? Why did everyone accept them until now? And how did television, arguably the most powerful influence on American culture, become so dependent on such a fragile collection of digits in the first place?

Sonar in the Living Room

THIS MAY BE HARD TO IMAGINE, BUT AT THE dawn of broadcasting, advertisers doubted that radio and TV would ever be useful for wooing consumers. They needed proof. They hired market researchers, themselves newly arrived on the commercial landscape, to conduct surveys, first simply to find out how many homes in America had a radio and later to determine what programs drew the most listeners. Somewhere along the way, however, ratings took on a power they were never meant to have, much to the chagrin even of Robert Elder, a co-inventor of Nielsen's first meter. In a 1978 letter he wrote that television “suffers greatly from the misuse of the [Nielsen Television Index] and for that reason I am not too happy about my part in getting it started.”

Nielsen, founded in 1923 by Arthur C. Nielsen, Sr., acquired the rights to the Audimeter from Elder and his partner, Louis Woodruff, in 1936. The company spent six years refining the idea, and in 1942 launched a radio ratings service that within a few years went national and directly challenged the dominant supplier of the day, C. E. Hooper, Inc., whose “Hooperatings” still relied on surveys. Nielsen promised better, faster numbers, and quickly left Hooper in the dust. In 1950 Nielsen launched its meter-based television index, all the while continuing a policy of aggressively patenting every conceivable audience-measurement idea—a policy that paid

off in 1961, when its toughest competitor yet, the American Research Bureau, which had crowded one of Nielsen's meter patents, wound up having to pay Nielsen royalties on all sales associated with the technology. The ARB temporarily abandoned TV meters in 1972; it took the name Arbitron in 1973.

Throughout the 1950s the amount of data that Nielsen delivered and the speed at which it did so increased sharply. At first Nielsen needed six weeks for delivery. By 1961 delivery took sixteen days, although in 1959 Nielsen had begun reporting day-after ratings from a small sample of homes (ARB did it first, in 1957). Newspapers reported Nielsen's rankings, a practice that strengthened the importance of ratings. C. E. Hooper had established the practice, by providing lists of top-rated programs to the press. The growing influence of ratings twice drew the attention of Congress, more forcefully in 1963–1964, with a full-blown investigation that yielded 1,700 pages of testimony and, ultimately, a set of Federal Trade Commission guidelines that sternly warned broadcasters always to remember that ratings data, based as they were on statistical estimates, were “inherently imperfect.”

Researchers, meanwhile, took up the hunt for a passive meter.

In 1964 a San Diego inventor named James Tanner caused a major stir at the National Association of Broadcasters' meeting in Chicago, where he planned to demonstrate a ratings truck capable of capturing the tuning of televisions that it passed on the street. The truck, he claimed, could monitor up to 3,500 TVs every thirty minutes while moving at speeds of up to fifty miles per hour, giving new meaning to the notion of a ratings sweep.

Tanner had to cancel his demonstration, however, because he couldn't get the truck to work. The idea died.

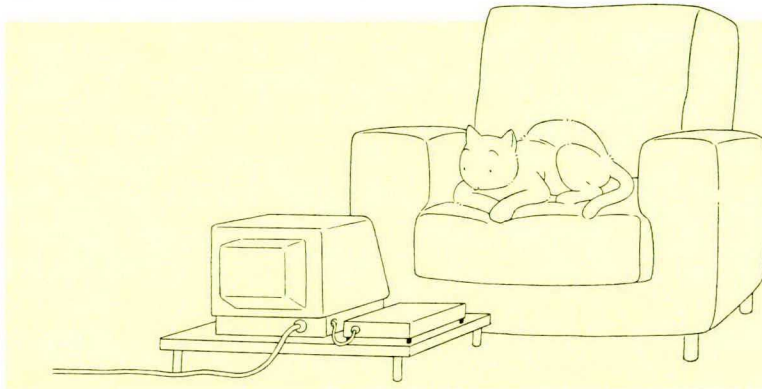
The next year an Oklahoma State University professor, Charles L. Allen, produced a sensation when he published the results of his work with what may have been the first passive audience meter, his DynaScope, a stop-action movie camera that, with the help of a strategically placed mirror, took pictures every fifteen seconds of both the audience and the shows they were watching. Graduate students analyzed almost 1.5 million photographs from ninety-five test homes to see whether people were attentive, inattentive, or simply absent while the TV was on, and discovered that 19 percent of the time the TV played to an empty room. For another 21 percent of the time someone was present but so involved in something else that he or she didn't see the screen. The DynaScope didn't restrain anyone's behavior, apparently: Allen's graduate students wound up poring over stop-action accounts of couples fighting and making love.

The study spurred a flurry of invention among Nielsen's engineers. They experimented with bladderlike scales hidden in couches which would identify people by weight, a concept later embodied in what employees still

TVision Insights, Inc. – Ex. 1016, Page 000003

refer to as the “whoopee cushion” patent. The scales, however, couldn’t tell the difference between kids and big dogs. The engineers also tried using height as an identifier. They lined doorways with electric eyes that measured each viewer who came through. High heels were its undoing.

Things took a turn toward the spooky in 1981, when David A. Kiewit became Nielsen Media’s director of engineering. He had spent the previous decade building surveillance and targeting systems for submarines and bombers, a terrific background for anyone interested in either counting or destroying vast numbers of TV viewers. Kiewit and colleagues devised an infrared sensor that scanned rooms for “hot bodies.” The system could indeed locate and count people, but it also counted light bulbs and toasters. One Nielsen engineer took the system home for some real-world testing and quickly got fed up. It kept reporting an extra person in his living room.



“Whenever anybody upstairs took a shower,” Kiewit told me recently, “there appeared to be a very long, thin person standing behind the couch who stood there for ten minutes or so and then *slo-o-o-owly* went away.” The phantom was a hot-water pipe embedded in the wall.

With a little tinkering the system could be calibrated to skip most nonhuman sources of heat, but it, too, proved vulnerable to the big-dog effect. “You could set your threshold so that if a house had a small dog, you could skip him,” Kiewit said, “but a big dog? You were in trouble.”

Next Kiewit’s group built a wall-mounted sonar set that scanned living rooms in exactly the way a ship scans New York harbor, transmitting inaudible signals and capturing the returning echoes. The system compared successive sonar images to isolate points of movement, and presumed these to be members of the audience. It was sensitive enough to pick up the soft heave of a viewer’s chest as he breathed—or the soft heave of a seat cushion easing back into shape after its occupant had gotten up and left.

As elaborate as these systems were, neither even came close to identifying viewers. A true passive meter was still a long way off, and Nielsen wasn’t in any great rush to get there. Its monopoly was intact; no serious competition was in sight.

The landscape of television, however, had suddenly

begun to change. In the 1950s a typical family got approximately three channels—well, actually 3.8 by Nielsen’s count. The networks sliced the audience pie into three substantial portions and didn’t much worry about whether Nielsen correctly allocated the crumbs. But cable TV arrived, VCRs followed (they first hit the U.S. market in 1975), and remote controls allowed viewers to switch channels effortlessly. The networks’ share of audience began to shrink.

A word of explanation here: Nielsen reports two numbers, the rating and the share. The rating is a measure of average audience and is expressed as a percentage of America’s total potential audience of 93.1 million TV-equipped households. During prime time only about 60 percent of these households actually watch their TVs. Nielsen’s share figures show how this active nightly audience divides itself among the various programs and stations. Suppose, for example, that this Thursday night 50

Nielsen began documenting the networks’ declining audience in the 1978–1979 season, when the three still drew a robust 91 percent of prime-time viewers.

million households watch TV. Suppose that 25 million of these active households watch *Cheers*. The episode’s rating would be 26.9 gross rating points, or 26.9 percent of all 93.1 million TV-equipped households. The share would be 50 percent, or half of the 50 million viewers who actually watched TV during that portion of the evening.

Nielsen began documenting the cruel fact of the networks’ declining audience in the 1978–1979 season, when the three still drew a robust 91 percent of prime-time television viewers. Their share has shrunk every year since. By 1985 the average household could get 18.8 channels, and the networks’ share of audience had shrunk to 77 percent. By last year the average home got thirty channels and the networks’ share had fallen to 63 percent.

For a bit of culture shock, consider this: Seven of the fifty highest-rated programs of all time are 1964 episodes of *The Beverly Hillbillies*. Each had a rating of 41.8 to 44. In the most heavily watched of these, Jed, Jethro, and Granny somehow managed to command a 65 percent share of viewers—more than the combined share of all three networks today.

By the mid-1980s the explosion in TV choices had begun to outpace Nielsen’s ability to keep up. By now advertisers cared less about reaching masses of households

TVision Insights, Inc. – Ex. 1016, Page 000004

and more about reaching the demographic segments of the population best suited to their products: for example, women twenty-five to fifty-four, men eighteen to forty-nine, or kids. Nielsen's set meters, sophisticated versions of the original Audimeter, recorded only that a TV was on and that it was receiving a particular station. Nielsen got its demographic data from viewer diaries kept by a second panel of families. Researchers feared that one member of each family, usually the already harried woman of the house, typically kept the diary for everyone. With so many channels, they worried, how could anyone remember exactly what a family watched, let alone the channel on which it appeared? They fretted, too, about the so-called halo effect that occurred when distracted diary-keepers wrote down what they *would* have watched or *should* have watched, rather than what they really did watch.

By 1985 the networks were ready for a change, and AGB, the British ratings giant, offered a tempting one. AGB had previously announced that it wanted to wire America with a new kind of meter, one it had already deployed in Ireland, Thailand, and the Philippines, which captured tuning and demographics simultaneously, continuously. AGB offered all this, moreover, at a lower cost.

The networks, delighted by the challenge to Nielsen's monopoly as well as by the promised improvements, helped fund an initial test of AGB's technology in Boston. This was a tricky time for Nielsen. Even though the networks had already lost market share, they were still Nielsen's biggest customers; on a symbolic level, moreover, their continued endorsement was vital. But their contracts with Nielsen were about to expire. Nielsen had offered each of them a new three-year contract that would cost \$14 million.

CBS made a show of canceling its contract, ostensibly to prevent automatic renewal; in a further display of disaffection it signed a one-year contract with AGB. In July of 1987 David Poltrack, then and now CBS's top research official, told *The New York Times*, "It is certainly possible that we will not be a client of Nielsen come the fall." ABC also canceled, but did not sign with AGB.

AGB's year-and-a-half-long Boston test gave Nielsen time to dust off a meter it had begun testing a decade before, which worked very much like AGB's fearsome new weapon. Both were called people meters, because they sought to provide an electronic measure not just of tuning but of who was watching. Both began producing ratings in September of 1987, and thus gave broadcasters an opportunity to see how competing ratings systems would describe the same objective reality.

The two new systems produced radically conflicting results. The variations, moreover, occurred in a pattern that defied reduction to a mathematical constant. On the first Monday night, for example, Nielsen's people meter rated ABC's football game 15.3. AGB gave it 11.9, a discrepancy of roughly three million households. Nielsen's meter-and-diary system, which the company continued

in order to provide stability during the transition, gave the game 14.5.

Nielsen flatly declared that AGB's numbers were "wrong." AGB said Nielsen's were too high. Confusion reigned; advertising negotiations faltered. Ultimately the industry voted to stay with Nielsen, which was at least a known entity. AGB fled the country, reportedly trailing \$80 million in losses.

The new meter brought more data more quickly, including overnight demographics, but it was no friend to the networks. That first season the networks' share of audience dropped from 75 percent to 70 percent. The networks, acting together as the Committee on Nationwide Television Audience Measurement (CONTAM), a body that had been founded in 1963 in response to the congressional hearings, hired a respected media-research firm, Statistical Research, Inc., of Westfield, New Jersey, to investigate the new system.

The resulting report was completed in late 1989 and weighed six pounds. It was 550 pages long, in seven volumes, and arrived in a slick, heavy-gauge gray box meant to last a century. I read the report—every killing, if soporific, detail. This may be my imagination, but I thought I detected a kind of deep and weary sorrow in its text, a disappointed sigh: so this is what went on behind the curtains of Oz all these years! The report "documented an erosion in the quality of the Nielsen service," in its authors' words. "The vision, the actions, and the commitment of the past have been diminished."

Jonn Dimling, of Nielsen Media, read the final, most critical volume in his room in a Northbrook, Illinois, Holiday Inn. "I was about as depressed as I've ever been," he told me. "Even though I was prepared for everything they were going to say, seeing it all, just one thing after another—you think, 'My God, what have we done?'"

To make matters worse, a few months later Nielsen reported the sharp declines in some audience segments which so troubled the networks. Viewers disappeared. In particular, kids failed to show up for Saturday-morning cartoons. Drove of viewers in key segments deserted the prime-time hours. The declines were so steep, so unexpected, that they caused an immediate combined loss to the networks of about \$200 million, the value of commercial time the networks had to cough up to make good on previously arranged audience guarantees.

In the months following, the networks launched an attack on Nielsen unprecedented for its ferocity. The assault, however, drew attention to what may have been the most significant revelation of the CONTAM report: how devilishly complex this business of measuring America's TV audience has become. Ronald Kessler, the program director of the prestigious Survey Research Center at the University of Michigan, and both a critic and a closet admirer of Nielsen, thinks that Nielsen's ability to produce any ratings at all is something of a minor miracle. "If you look inside what Nielsen does, it is absolutely

TVision Insights, Inc. – Ex. 1016, Page 000005

mind-boggling," he says. "Push any door and you find unbelievable detail."

At the same time, he added, "little errors mount up like crazy in a thing like this."

Choosing Nielsen Families

I DON'T KNOW ANY NIELSEN FAMILIES. I HAVE NEVER met a member of a Nielsen family, and to my knowledge no Nielsen families live on my block. The only person I know who even comes close to having known a Nielsen family is my wife, who recalls vague feelings of childhood envy toward a family in her neighborhood which had been recruited to the cause. At any one time about 4,000 American families are wired by Nielsen, or one for every 23,000 of the country's television-equipped households. They are a secretive bunch, under orders to keep quiet about their Nielsen status lest someone try to coerce them into a scheme to manipulate the ratings. (They are told to report any suspicious encounters.) Together they compose one of the most influential groups in America, for their viewing habits directly and immediately determine the allocation of roughly \$10 billion worth of network air time each year. A single Nielsen household, therefore, accounts for about \$2.5 million of that \$10 billion in advertising fees. Ronald Kessler puts it another way: "It's worth a television network's time to find a \$150,000 house in America that has a Nielsen meter, pay the owner \$200,000 for the house, and then tune that television set to their station a hundred percent of the time."

Broadcasters treat Nielsen's ratings as if they were exact numbers, not estimates. The networks sell about 70 percent of their prime-time advertising in what is called the up-front market, a time bazaar that opens in June, shortly after the networks announce their schedules for the following fall. Here advertisers buy big blocks of time in advance; in return the networks sell the time at a discount and guarantee that the advertisers' commercials will draw an agreed-on number of viewers. Deals are cut to the tenth of a ratings point; viewers become a commodity, sold on a cost-per-thousand (CPM) basis. An unexpected ratings dip of a single percentage point can cost a network millions of dollars in "make-goods" to satisfy the guarantees.

"As negotiations take place, deals are decided on the basis of CPM differences of pennies," CBS's David Poltrack said in a recent talk to members of the Advertising Research Foundation. Yet these negotiations, he said, ignore the range of error inherent in any statistical sample, and the far greater error in a sample as fundamentally flawed as Nielsen's. A true calculation of the room for error, he said, would "tell us that even a CPM variation of a dollar or more may not be real."

Statistical theory holds that every sample, no matter how perfect—even if everyone surveyed agreed to par-

ticipate—starts life with a built-in, or standard, error. If a single episode of, say, *Murphy Brown* earns a rating of 15, the standard error is six tenths of a ratings point in either direction, according to Edward Schillmoeller, Nielsen Media's chief statistician. The episode has roughly a 68 percent chance of having an audience size between 15.6 and 14.4, and a 32 percent chance of having an audience that falls outside those limits.

This level of accuracy, however, applies only in ideal, textbook conditions, not in the rough-and-tumble world of surveying people, where the degree of error may be far greater.

Nonetheless, small changes in ratings affect even programming decisions. Poltrack told his audience, "Programs are canceled or shifted around the schedule after a four-week ratings trend is provided, even though differences between the points in the trend line may not be significant. In short, we have ignored the weak foundation of our television measurement system. . . ."

Daily ratings intensify this false sense of precision. "We could probably do without daily ratings," Poltrack told me during an interview in his office last April. They tend, he said, to shorten the already short-range perspective of his industry. The networks, faced with shrinking audiences and rising costs, have little patience. "*Murphy Brown* took three years to be a hit," he said. "This is a mentality where if you're not a hit in four weeks, we take you off and put something else on."

Having said this, however, he tipped back his chair and began flipping through a set of Nielsen overnights that had just arrived on his desk. "There's always *something* interesting," he said. He turned to the ratings for the previous Friday and cited that night's episode of the CBS series *Dallas*. "*Dallas* has been going up in the ratings for us, and *Dallas* is going to have its last telecast ever in a couple of weeks. We want to see if there's any momentum building. And sure enough"—he pointed to the page—" *Dallas* is up again in the ratings, a good sign there's some momentum."

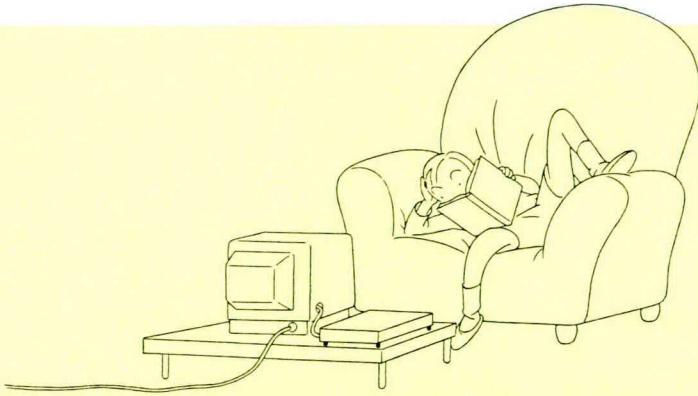
The news about another show was less sanguine. He turned to Thursday's revised ratings. CBS had recently introduced a lackluster stopgap series called *The Antagonists* opposite NBC's top-rated comic juggernaut, *Cheers*.

"The first couple of weeks it got okay ratings," Poltrack said. "We were waiting very anxiously to see what the show did when *Cheers* went into repeats."

That previous Thursday a fresh *Antagonists* had aired against the first *Cheers* repeat of the season. Nonetheless, *Cheers* got an 18.8 rating, or a 31 share; *The Antagonists* got only an 8.1 rating, for a 13 share.

"It didn't go up at all," he said. "That's bad news, because the rating was not high enough to represent success for the show, and the fact it didn't go up at all . . ." He stopped, and let the thought—and the show, as it happens—hang. *The Antagonists* didn't make it to this season's schedule.

TVision Insights, Inc. – Ex. 1016, Page 000006



By tracking the audience on a second-by-second basis, Nielsen's passive meter may yield the first accurate ratings of commercials—not shows, commercials.

Ratings have attained a certain psychic precision as well. A fraction of a point can mean the difference between a producer's being considered a winner or a loser, and in television the distinction matters. "It's not the reality, it's the perception," says Bonny Dore, a successful producer of TV movies and miniseries, such as *Sins* and *Glory! Glory!* "If the perception is you lost—you lost," she says. And that perception can affect negotiations for a producer's next deal—"or whether you *get* another deal."

"Does it bother you," I asked, "that Nielsen's ratings are so flawed?"

"You're simply feeding into every nightmare a producer's ever had. We work and slave and go into debt and we do something wonderful, and the numbers are, quote, 'a little disappointing.' Then to find out that maybe they're off by five share points? You just want to go kill yourself. Other than that," she said, laughing, "it's fine."

Where Statistical Sampling Goes Wrong

AT ITS VERY BEST, STATISTICAL SAMPLING IS fraught with problems. The initial response rate for the 1990 census was the lowest in history, even though all we had to do was fill out a single questionnaire, and are required by law to do it. Nielsen runs a huge, continuous, partly automated survey of America twenty-four hours a day, 365 days a year, and intrudes on the daily life of each recruit for two years at a stretch.

In the real world, as Nielsen has found, sampling is hell.

The process begins with the creation every ten years of a "sample frame," the population of households in America from which all Nielsen families are selected. First, using Census Bureau data, Edward Schillmoeller's department divides the country into successively smaller hunks of territory until at last it has a collection of some 5,000 "block groups," or units of census geography comprising on average about 425 housing units each. Next, Nielsen sends squads of field agents into each of these block groups to count every housing unit, both to account for any changes in the number of homes since the Census Bureau made its report and to note the addresses of po-

tential Nielsen candidates. The crews merely count, until they reach a number selected at random by Schillmoeller's department. Assume there are indeed 425 homes in a block group. Nielsen picks a random number from 1 to 425. If the number is 125, the enumerator counts the first 124 homes without jotting down a single address. The next housing unit, be it a cave, tree, or converted beer truck, is the primary target household, a "Basic" in Nielsen-speak. This is the home Nielsen most dearly wants to recruit, and the one that sampling theory dictates Nielsen indeed *must* recruit. If the house is vacant, or even a vacant slab in a trailer park, Nielsen will keep coming back for five years in hopes that someone will have moved in.

In theory, every household has an equal shot at becoming a Nielsen family; in practice, however, the odds for some of us are zero. Nielsen would never allow me to be a member of its sample, because I am "occupationally disqualified," as are all my fellow journalists, every editor, producer, and advertising executive, and of course every single employee of CBS, NBC, ABC, CNN, MTV, local television stations, and Arbitron. The presumption is that we could not be trusted to behave ourselves and at the first opportunity would rush out and write about the experience or tune our TVs to the stations we work for. (Nielsen has good reason to worry: station employees included by chance in diary panels have been known to "load up" the diaries with their own stations' programs.) Nielsen also builds in demographic safety valves to avoid the unlikely but statistically possible selection of an aberrant sample—say, 4,000 black households. Such an occurrence might delight the networks, given that black families watch half again as much television as their nonblack peers, but would not do much for Nielsen's statistical credibility.

Today's people meter demands a lot from its hosts. Arthur Nielsen, Sr., wrote in 1955 that one of the beauties of the Audimeter system was that it operated "entirely automatically and outside the collaborating family's day-to-day awareness." Today's meter is a nagging shrew, consisting of three components: a computer the size of a compact-disc player, known as the home unit; a smaller box with a numerical keypad and a faceplate of TVision Insights, Inc. — Ex. 1016, Page 000007

red and green lights, which sits atop the TV set; and a remote control, typically with eight numbered buttons and an "OK" button. Nielsen assigns a button to everyone in a household, even children as young as two. And everyone, even the two-year-old, is expected to behave in appropriate Nielsen fashion—that is, to press his button whenever he starts to watch TV and whenever he stops. Visitors pick an unoccupied button, enter their sex and age on the numerical keypad, and then also report their arrivals and departures. When a viewer punches in, the lights associated with his number change from red to green. Periodically the meter demands a little reassurance that someone is indeed watching, and will blink a single red bulb at the far right of the set-top box. One press of the "OK" button calms the machine. The longer the prompt goes unanswered, however, the more adamant the meter becomes, blinking its bulbs in an increasingly frenzied manner until it is ablaze with flashing lights, like the scene of a midnight car wreck. Some meters finally cry out in frustration, emitting a pained and disappointed beep.

Volunteers also subject their homes to a lot of drilling and wiring. Nielsen requires that every single TV set, VCR, and cable converter be metered and connected to the home unit. Any satellite dish has to be metered separately with an "inclinometer" that records which way, and thus at which satellite, the dish is pointing. Nielsen even takes an inventory of unused TVs kept in the attic or basement. Nielsen's agents don't simply accept a homeowner's assurance that a set won't be used. They have him swear to it in an affidavit, and then attach a small plastic seal to the prongs of the plug. If on a return visit the agent finds the seal broken, he will ask permission to meter the set; if he isn't allowed to, he may eject the family from the sample.

Today's typical Nielsen home, says Larry Patterson, the company's director of field operations, has two TVs, one VCR, and one cable converter, requiring an array of equipment that takes five or six hours to install. The most complex house he has come across had six TVs, five VCRs, and cable access; installing the meters took two full days.

In return for all this equipping and pestering, each family can pick gifts from a Nielsen catalogue and gets cash payments of two dollars a month for each TV or VCR in the house. A household with one TV would get just \$24 a year; the family with the complex system mentioned above would get \$264.

The networks worry that the burden is too great, that Nielsen families experience "button fatigue" and begin engaging in such inappropriate behaviors as failing to log in, or letting a visitor *just sit there* without first registering his age and presence. CONTAM conducted exit interviews with families leaving the Nielsen stable and found that one out of three reported getting tired of the meter; half of these said that as a result accuracy had suffered. More

than one in ten said they "hardly ever" logged out when they finished watching a show.

The networks fear that button fatigue is especially pronounced among children. Almost half the kids (or their parental spokespeople) interviewed by CONTAM said that they did indeed get tired of pushing buttons. Nielsen takes great pains to win their cooperation. It gives them a Nielsen coloring book and shows them a video called "The Nielsen People Meter Show." Nielsen wants kids to bond to their buttons, and practices some subtle human engineering. It gives them a choice of animal stickers to place next to their buttons—but not commercial figures like Donald Duck or the Ninja Turtles, of course, because these might bias their viewing behavior.

Nielsen assigns the youngest child button No. 1 and works up from there by age. It does this to avoid button "churn," the shuffling of meter slots as people join or leave the household. Nielsen wants the oldest person on the highest button because of the cold actuarial fact that this person is the likeliest to die and leave a disorienting gap in the panel.

Who would agree to invite such an imperious guest into their homes? Far too few people, as it happens. In a typical week only about 48 percent of the households installed and on line are Basics, according to CONTAM. Such a low rate of cooperation makes statisticians queasy; it could be an indicator of a fundamental difference between those who do cooperate and those who don't. The question is, how does this affect viewing patterns? An early CONTAM study, conducted in 1964 in response to the congressional investigations, found that those who cooperated watched somewhat more TV than non-cooperators. The networks claim that the people meter has reversed the pattern—that now people who don't cooperate are the heaviest network watchers.

Nielsen's cooperation rate is actually far worse than 48 percent. In its routine disclosures the company reports only the percentage of Basics who refuse to participate, and doesn't mention the number of "Alternates" it had to approach before achieving its 4,000-household sample. In fact, Nielsen must approach three Alternates, on average, before it finds a willing participant. In some cases, according to Schillmoeller, recruiters have had to approach up to thirty additional homes.

Moreover, on any given day about 10 percent of Nielsen's people-meter households produce faulty data that cannot be included in the daily ratings calculations. The families most likely to be excluded, CONTAM discovered, were the heaviest viewers, reflecting a Catch-22 phenomenon literally engineered into the people-meter system: the more TVs and VCRs and Nielsen equipment you have, the greater the chance something will happen to bump your data from the day's tabulations.

Fatigue causes some families to drop out of the sample altogether. Other households leave for quirkier reasons: TVision Insights, Inc. – Ex. 1016, Page 000008

one family bought a slick new TV and didn't want Nielsen's installers tinkering with it. Larry Patterson lost another Nielsen household to one of the cold realities of ghetto life in New York City. He and a partner had recruited an elderly woman living in a tough housing project in the borough of Queens. Early the following week he got a call from a city homicide detective who told him that the woman had been found strangled in her bed the previous Sunday morning. He requested all the woman's viewing records for Saturday night, in case they held some clue to the timing of her murder.

She had turned off the set at 11:23 P.M. on Saturday, Patterson found. But he also discovered a curious thing: someone had watched a lot of television in the woman's apartment the Sunday afternoon after she'd been killed.

The detective was silent a moment. "Oh, yeah," he said. "That was us. We were watching the football game."

"For him, at least," Patterson told me, "it certainly confirmed our credibility."

Readings from America's 4,000 people-metered households are automatically transmitted every night from 3:00 A.M. to about 7:00 A.M. Florida time, the machines dutifully calling home to Nielsen's Dunedin Center. Most people assume that the people meter alone produces all the information Nielsen needs for its ratings. In fact this nightly torrent of data is by itself utterly useless.

The meter tells only who is watching the TV set and how the set is tuned. Nielsen must merge this data with a second torrent of information, the detailed records of exactly what programs appeared on thousands upon thousands of TV stations and cable "headends" around the country—not what the stations planned to broadcast but what they actually did broadcast, and the exact times at which they did so.

On any day the task is monumental: The networks supply Nielsen with their daily lineups, but typically these succumb to special reports, late-running baseball games, natural disasters, and so forth—regional events that might cause an affiliate to interrupt the network "feed." Nielsen keeps track of all this by stationing robotic TV watchers in the TV markets where Nielsen families reside to watch television twenty-four hours a day, *forever*. They watch for a basic network signal, hidden from human viewers, that indicates when the network feed is being broadcast by an affiliate. If the signal arrives at its scheduled time, Nielsen assumes that all is well. Any discrepancy is called a conflict. In a typical week Nielsen's program-records department in Dunedin will resolve 12,000 to 15,000 such conflicts, often by calling stations to find out just exactly what did air. During the first week of the Gulf War, Nielsen confronted 50,000 conflicts and in the process exposed a bit of affiliate treachery. A few stations had dumped their network feeds to broadcast CNN instead.

CONTAM factored together all the forces tending to re-

duce cooperation and the supply of usable data, and determined that Nielsen's people-meter sample had a final net response rate of only 37 percent, meaning that only 37 percent of Nielsen's originally targeted Basics produced usable household data on any given day—a rate "below all industry or government standards for high quality operations," CONTAM reported. "The rate has declined steadily over recent years and indications are that, unless a different effort is introduced, it will continue to do so. At some point . . . a totally different measurement approach will have to be considered."

The New Passive People Meter

"**A**S YOU CAN SEE, THIS IS YOUR TYPICAL LIVING room," said Christopher Lubniewski, a Nielsen client-service representative, in a recent interview. The room was a barren, windowless office in Nielsen Media's executive headquarters in New York. Nielsen's new passive people meter, a squat black box the size of a videocassette recorder, was perched before me on the shelf of an incongruous breakfront, the only thing even remotely living-room-like.

The machine and I had already been introduced—that is, it had scanned me and then stored four images of my face: three head-on portraits and one quarter-profile. It was now ready to try matching the real me to these digital mug shots.

If it succeeded, it would display on a nearby video screen my name and a number indicating how confident it was about the identification. A number and no name would mean that it had failed. I did not want a zero. A zero would be a pretty good indicator that I had died or at least stopped breathing. A number from 60 to 83 would mean that I was a person, although not someone the meter knew. Anything higher would mean that it knew me well enough to display my name.

The meter scans the room every two seconds, looking for signs of life. It compares each fresh image in a given location with its predecessor to produce a digital representation of movement, the hallmark of humanness. A life-size photograph brought into the room might fool the system for a moment, Lubniewski told me, but over the next few scans its lack of movement would cause the meter to ignore it.

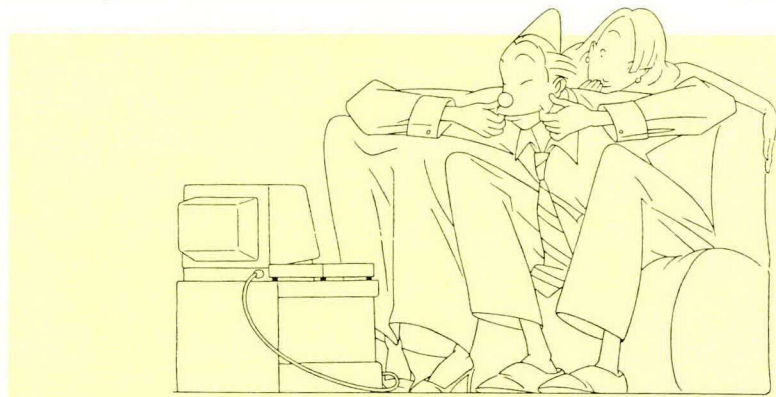
The system next locates the melange of digits most resembling those for a human head, a search program that may at last help Nielsen separate the men from the dogs. "If a cat or a big dog jumped up here," Lubniewski said, "the system would scan it and try to find the head, and during the match-off would give it a score of zero to sixty." A nonhuman moving object.

The machine notices every departure and return. Its reliance on head-on or nearly head-on views, moreover, means that it will not count people who are present but TVision Insights, Inc. — Ex. 1016, Page 000009

involved in other activities, such as reading the newspaper or making love on the floor.

The machine gets good but deeply qualified reviews.

"It's very accurate," Robert A. Warrens, a senior vice-president and the director of media resources and research for J. Walter Thompson USA, told me not long ago. "I sat in a room in different positions and overall it was able to pick me up a hundred percent." But many who have seen the meter worry that it could worsen Nielsen's already poor cooperation rate, especially given today's heightened concern for privacy. "We know that one out of every two households has a TV in one or more bedrooms," says Nicholas P. Schiavone, NBC's vice-president of media and marketing research, and the current chairman of CON-TAM. "Now, I come to your home and tell you I'm going to set up a recognition meter in every room of your household where there's a TV. Will people allow a device that can recognize—that can *see*—into those rooms?"



"It scans the room every one or two seconds. You could fool it, but chances are if it misses you—if you sneeze or make a face—we'll still know you're there."

Nielsen installed a nonfunctioning version in a group of Nielsen homes to see if their occupants, led to believe the system really worked, would indeed accept a camera-like device in their living rooms, bedrooms, even bathrooms (where one percent of American families currently have a TV). The answer, apparently, is a qualified yes.

The thing could appear in viewers' homes fairly soon. William G. Jacobi, Nielsen Media's president and chief operating officer, told me he'd "be disappointed if the passive meter were not playing a role in the rating system within the next three years."

If the meter works, it will solve Nielsen's problem of button fatigue. One school of thought holds that the passive meter could even improve initial cooperation rates by eliminating the workload associated with the current people meter. It would do nothing, however, to help Nielsen keep up with the proliferation of miniature portable TVs and track the ranks of so-called "out-of-home" viewers, the great uncounted masses who watch television while on vacation, in hotels, away at college, or on the job. To help meet these challenges, Nielsen's Dunedin researchers are experimenting with a pocket-sized electronic "diary" that people would, in theory, whip out to record the shows they watch in bars and elsewhere. They are also testing ways by which various

pieces of Nielsen equipment could communicate with one another through the ordinary plugs and circuits of a house, thereby eliminating the need to run so many wires through homes. Moreover, Nielsen recently hired its first chief research officer, a kind of research cop, to study Nielsen's existing rating methods and look for simple ways to improve them.

The passive meter, however, is the thing Nielsen most likes to show off. "Try moving your chair backward and forward," Chris Lubniewski suggested. The meter registered my identity with scores of about 90.

Lubniewski took a seat in an adjacent wheeled chair and together we kicked our way throughout the meter's field of vision. It named us both. I tried covering the top half of my head with a legal pad, something I often do when I watch TV. The machine hesitated. I was a person, all right, but it couldn't say who.

Changes in hairstyle won't throw the meter, Lub-

niewski said—although he did mention a recent exception. "One woman came in here wearing her hair very tight," he said. "She let it go and it fell in front of her face, and the system *did* get confused."

This could be forgiven, I decided. In a similar situation I'd have felt likewise—but what about shaving a moustache? No effect, Lubniewski said; someone had tried it.

I asked Nielsen's public-relations man, sitting nearby, if I could borrow his glasses. The machine named me with its highest score yet.

It was time to get serious. Man to machine. I stuck out my tongue, yanked back the corners of my mouth, and closed one eye, a maneuver I picked up from my three-year-old daughter.

Lubniewski smiled, but the smile looked thin to me; it masked, I was certain, the kind of ennui that can come only from watching armies of otherwise sober people suddenly let down their hair or flourish surprise pairs of glasses.

"A lot of people try to fool the system," Lubniewski said. "You have to remember, it scans the room every one or two seconds. You could fool it, but chances are if it misses you—if you sneeze or make a face—we'll still know you're there."

TVision Insights, Inc. – Ex. 1016, Page 000010

I smiled past my fingers but held the pose a bit longer, just to make sure the system got it. It identified me as human, but otherwise didn't recognize me.

Satisfied, I tucked in my tongue.

An instant later, the monitor displayed my name with a cautious score of 85.

What Lies Ahead

A TRUE PASSIVE METER IS ALMOST CERTAIN TO APPEAR over the next few years. Barring the unexpected arrival of a new challenger, that meter will function like the one I saw demonstrated in Nielsen's offices. In the past, fundamental advances in ratings technology have brought far-reaching change to the business and culture of television. What changes will this meter bring? What revelations about the audience-holding power of commercials? Of TV itself?

History is full of broad hints about the awful truth: plunging water pressure in Toledo, the DynaScopic chronicles of TV sex in Oklahoma. The best indications of what lies ahead, however, come from the deployment of near-passive meters here and in France. In 1987 R. D. Percy & Company, a now-defunct ratings concern, installed an infrared body sensor in 609 homes, and demonstrated that television sets played to empty rooms about 13 percent of the time. The proportion climbed sharply for certain categories of shows. Golf commentators, for example, spent 20 percent of their time whispering stroke-by-strokes to unoccupied couches and chairs. Viewers abandoned *Good Morning America* and *Today* for roughly twelve minutes out of every hour. During commercials the absentee rate as much as doubled.

French advertisers recently got a look at commercial avoidance. In France, Mediametrie holds a virtual monopoly over TV ratings. In 1989, however, a challenger, Telemetric, deployed a form of passive meter called Motivac, which detects the presence of people by sensing motion, just as Nielsen's does. Its ability to identify viewers, however, depends on the fact that the French watch TV the way Americans did back in the 1960s—that is, with the whole family gathered in favorite seats around a single TV. Viewers must provide Motivac with detailed information about their TV-watching habits, including where they most like to sit. If Motivac finds a person in that *chaise favorite*, it checks its memory for the likely occupant.

Motivac provides Telemetric's clients with a moment-by-moment "spike graph" of the nation's viewing; it proved with dismaying immediacy that vast numbers of viewers flee prime-time shows during commercials and the final credits. It found, for example, that a collection of commercials broadcast during a World Cup soccer match triggered a virtual stampede from the living rooms of France, with the TV audience suddenly shrinking from 9 million to 3.5 million. Telemetric also discovered that

high ratings garnered by French news shows during the Gulf War were deceptive. Viewers gathered only at the beginning of each broadcast, to catch the headlines, and then left the TV on and went about their business elsewhere in the home.

Such news here could cause an immediate restructuring of the pricing of advertising time. No advertiser believes that everyone stays for the commercials, of course, but a second-by-second measure of audience size would at long last document the rate of desertion. Advertisers would discover that some shows hold their audiences more tenaciously during commercial breaks than others. For most of my friends, the Super Bowl is a big beer-and-popcorn event, but the half-life of beer in the human body is known to be brief. Suppose advertisers discover that most of America flees to the bathroom during those ultra-expensive bowl spots?

Commercial ratings might bring an element of justice to television. At last advertising agencies would find themselves subject to the same ratings tyranny that producers and TV writers have had to endure all these years. Suddenly agencies, too, would find their Nielsen rank published in *Variety* and *The Los Angeles Times*. Think of the implications—now *commercials* could be canceled!

Second-by-second ratings, however, could lash the creative content of shows even more tightly to the numbers. The new ratings would provide programming executives with an unprecedented ability to tinker with the content and structure of television shows. "It is conceivable," John Dimling told me recently, "that if you literally had second-by-second or minute-by-minute ratings, you'd know what people in a given demographic group tend to leave the program at what point in time. The use of that information in making programs could be kind of mind-boggling. You could do a content analysis of the program, and relate the content to when people tend to leave, or when they come back. Or if they come back. It may turn out that a certain point in the story or a certain production technique tends to interest people or not interest people."

Suppressing a chill, I asked, "You mean one could rate individual characters? Even subplots?"

"Sure, sure," he said, without an instant's hesitation.

If Nielsen now stands on a threshold, so does television itself. Ratings began simply as a way for advertisers to prove to themselves that radio and television could help them reach masses of consumers. The numbers soon took on a life all their own and an authority they were never meant to have. They became the soul of television. But when is enough enough? How much information is too much information? Does TV now risk snuffing the last ember of creativity under a yet huger cascade of deceptively precise numbers? Before Nielsen takes that next great technological leap, I'd like to invoke the old FTC warning: Even at their best, Nielsen's ratings will always be estimates. They will always be "inherently imperfect." □

TVision Insights, Inc. – Ex. 1016, Page 000011

Copyright of Atlantic Magazine Archive is the property of Atlantic Monthly Group LLC and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.