GOOD COMPUTING: A VIRTUE APPROACH TO COMPUTER ETHICS

A book under contract with
Jones & Bartlett Publishers
40 Tall Pine Drive
Sudbury, MA 01776
http://www.jbpub.com/

DRAFT FOR
CS 263
June Puerto Rico writing retreat

Chapter 6
Safety: Therac 25 and Hughes Aircraft
Version 1: 01/30/05 by Chuck Huff
Version 1.1: 03/2/05 by Chuck Huff
Version 1.2 8/1/05 by Chuck Huff
(revisions as agreed on at retreat)

Based on writings in www.computingcases.org prepared by Chuck Huff and Bill Frey and class documents from University of Mayaguez, PR, FILO 3185 prepared by Bill Frey

Therac-25

Table of Contents.

Abstract.

Therac-25 was a new generation medical linear accelerator introduced in 1983 for treating cancer. It incorporated the most recent computer control equipment. Therac-25's computerization made the laborious process of machine setup much easier for operators, and thus allowed them to spend minimal time in setting up the equipment. In addition to making setup easier, the computer also monitored the machine for safety. With the advent of computer control, hardware based safety mechanisms were transferred to the software. Hospitals were told that the Therac-25 medical linear accelerator had "so many safety mechanisms" that it was "virtually impossible" to overdose a patient. As it turned out, the computerized safety monitoring was not sufficient, and there were six documented cases of significant overdoses of radiation, three resulting in severe injury and three resulting in death from radiation sickness.

Historical Narrative.

Therac-25 was released on the market in 1983 by Atomic Energy Canada, Limited (AECL). In 1987, treatments with the eleven machines then in operation was suspended. Those machines were refitted with the safety devices required by the FDA and remained in service. No more accidents were reported from these machines. At about that time, the division of AECL that designed and manufactured Therac-25 became an independent company and changed its name.

The major innovations of Therac-25 were the double pass accelerator (allowing a more powerful accelerator to be fitted into a small space, at less cost) and the move to more complete computer control. The move to computer control allowed operators to set up the machine more quickly, giving them more time to speak with patients and making it possible to treat more patients in a day. Along with the move to computer control, most of the safety checks for the operation of the machine were moved to software and hardware safety interlocks removed.

•Early Therac Machines

The story of Therac-25 begins in the early 1970's when Atomic Energy Canada Limited (AECL) joined forces with a French company, CGR, to design and build a medical linear accelerator based on earlier CGR machines. The companies cooperated on the design and manufacture of two successful medical linear accelerators, the Therac-6 and its successor, the Therac-20. Both these machines were based on CGR designs that did not use computer control. The new machines added computer control, in addition to other innovations. The Therac-6 was the initial product of their collaboration and was designed to produce X-rays for radiation therapy. The Therac-20 was a much more powerful and versatile machine. It could produce two different kinds of radiation beams for treatment of deep and shallow tissue. AECL also produced other medical linear accelerators, including the Therac-4, a single mode electron beam machine.

•Development of Therac-25

In the early 1980's, AECL developed a much more space-efficient medical linear accelerator that was just as powerful and versatile as the Therac-20. Linear accelerators are more powerful the longer they are, and AECL found a way to fold the long beam-producing mechanism for a 25 MeV machine into a smaller space. In addition, this new version was somewhat less expensive to produce, since it used a less expensive beam production device (a magnetron instead of a klystron).

Finally, AECL intended to take advantage of increasing capability of computer software to make the machine easier to operate. The new Therac-25 was the result of a convergence of the new beam-folding technology with the ease of computer control, bringing with it the bonus of lower production costs. In addition to lower production costs, the computer control allowed faster setup of the machine for each patient. This meant that more patients could be treated in one day than with non-computerized linear accelerators.

The Therac-25's ancestors, Therac-20 and Therac-6, had used a minicomputer (a DEC PDP-11) to add some convenience to the standard hardware of a medical linear accelerator. They both could work without computer control. AECL determined to make its new model, Therac-25, a tightly-coupled combination of software and hardware. By this time, its collaboration with CGR had grown stale and AECL was bringing in its new beam folding technology (and the new Therac-25) on its own.

In tightly coupling the software and the hardware, AECL could use the software to monitor the state of the machine for proper operation and for safety. Previous versions, with designs based in models that predated computer control, had included independent circuits to monitor beam scanning and had mechanical interlocks to ensure the machine could not enter a state in which it could harm a patient. But with increased computer control, AECL decided not to duplicate this safety equipment in the Therac-25 (with additional cost savings), and to rely on software for policing these safety issues.

•Therac-25 goes to Market

In late 1982, Therac-25 was first offered to hospitals in a commercial version. It was eventually adopted by eleven institutions, six in Canada and five in the US. These included sites in Georgia, Texas, Washington State, and Hamilton, Ontario.

•Safety Analysis of Therac-25

In 1983, just after AECL made the Therac-25 commercially available, AECL performed a safety analysis of the machine using Fault Tree Analysis. This involved calculating the probabilities of the occurrence of varying hazards (e.g. an overdose) by specifying which causes of the hazard *must jointly occur* in order to produce the hazard.

Since much of the software had been taken from the Therac-6 and Therac-20 systems, and since these software systems had been running many years without detectable errors, the analysts assumed there were no design problems in the software. The analysts did consider software failures like "computer selects wrong mode" but assigned them probabilities like 4 x 10**-9. These sorts of probabilities are likely assigned based on the remote possibility of random errors produced by things like electromagnetic noise. They do not take into account the possibility of design flaws in the software.

•The accident history of Therac-25.

July 26, 1985: Francis HilI. In July of 1985, AECL was notified that a patient in Hamilton, Ontario had been overdosed. She was a 40-year old cancer patient at the Ontario Cancer Foundation clinic in Hamilton, in for her 24th Therac treatment for carcinoma of the cervix.

The Therac-25 operator activated the machine, but after 5 seconds, the Therac-25 shut down and showed an "H-tilt" error message. The computer screen indicated that no dose had been given, so the operator hit the "P" key for the "proceed" command. The Therac shut down in the same manner as before, reading "no dose," so the operator repeated the process a total of four times after the initial try.

After the fifth try, a hospital service technician was called but found no problems with the machine. Francis Hill left the clinic and the Therac was used with six other patients that day without any incidents. However, despite the fact that the Therac had indicated that no radiation dose had been given during Francis Hill's five therapy attempts that day, Hill complained of a burning sensation she described as an "electric tingling shock" in the treated area of her hip.

Hill returned for treatment three days later, on July 29, and was hospitalized for suspected radiation overexposure. She had considerable burning, pain and swelling in the treatment region of her hip.

The Hamilton clinic took the Therac-25 machine out of service and informed AECL of the incident. This was the first time AECL had heard from a clinic about an overdose problem with the Therac-25 machine. AECL sent a service engineer to investigate.

AECL reported to a range of stakeholders that there was a problem with the operation of Therac 25. The FDA, the Canadian Radiation Protection Board (the parallel Canadian agency to the FDA), and other Therac-25 users were all notified. Users were instructed to visually confirm that the Therac turntable was in the correct position for each use.

Because of the Hamilton accident, AECL issued a voluntary recall of the Therac-25 machines and the FDA audited AECL's modifications to the Therac. AECL could not reproduce the malfunction that had occurred but suspected some hardware errors in a switch that monitored the turntable position. A failure of this switch could result in the turntable being incorrectly positioned, and an unmodified electron beam striking the patient. The company redesigned the mechanism used to lock the turntable into place, redesigned the switch to detect position and it accompanying software. They then reported in November 1985 that this redesign was complete and that, given their safety analyses, the machine was now at least 10,000 times safer than before.

Francis Hill died on November 3, 1985 from cancer. An autopsy revealed that had the cancer not killed Hill, a total hip replacement would have been necessary because of the radiation overexposure.

November, 1985: Katy Yarborough. In November of 1985, AECL heard of another incident in Georgia. On June 3, 1985, 61-year old Katy Yarbrough had been receiving follow-up treatment at the Kennestone Regional Oncology Center (Marietta, GA) for the removal of a malignant breast tumor. On June 3, staff at Kennestone prepared Yarbrough for electron treatment to the clavicle area, using the Therac-25 machine.

Yarbrough had been through the process before, which was ordinarily uneventful. This time, when the machine was turned on, Yarbrough felt a "tremendous force of heat… this red-hot sensation." When the technician re-entered the therapy room, Yarbrough said, "you burned me." The technician replied that that was "not possible."

Back home, the skin above Yarbrough's left breast began swelling. The pain was so great that she checked in at Atlanta's West Paces Ferry Hospital a few days after the Therac incident. For a week, doctors at West Paces Ferry continued to send Yarbrough back to Kennestone for Therac treatment, but when the welt on her chest began to break

down and lose layers of skin, Yarbrough refused to undergo any more radiation treatment.

About two weeks later, the physicist at Kennestone noticed that Yarbrough had a matching burn on her back, as though the burn had gone through her body. The swelling on her back had also begun to slough off skin. Yarbrough was in great pain, and her shoulder had become immobile. These clues led the physicist to conclude that Yarbrough had indeed suffered a major radiation burn. Yarbrough had probably received one or two radiation doses in the 20,000-rad (radiation absorbed dose) range, well above the typical prescribed dosage of around 200-rads. The physicist called AECL and, without telling of the accident, asked questions about the likelihood of radiation overexposure from the Therac 25 machine: Could Therac 25 operate in electron mode without scanning to spread the beam? Three days later AECL engineers called back to say this was not possible.

Katy Yarbrough was in constant pain, lost the use of her shoulder and arm, and her left breast had to be removed because of the radiation burns. Ms. Yarbrough) filed suit in November, 1985. There is no evidence that AECL followed up this case with the Georgia hospital. Though this information was clearly received by AECL, there is no evidence that this information was communicated internally to engineers or others who responded to later accidents. This lack of internal communication is likely the cause of later statements that there was no history of overdosing with the machine.

In January of 1986, AECL heard from a hospital in Yakima, Washington that a patient had been overdosed. The AECL technical support supervisor spoke with the Yakima hospital staff on the phone, and contacted them by letter indicating that he did not think the damage they reported was caused by the Therac-25 machine. He also notified them that there have "apparently been no other instances of similar damage to this or other patients."

December, 1985: Undisclosed name. The individual was being treated with the Therac-25 machine at the Yakima Valley Memorial Hospital in Yakima, Washington. After one treatment in December 1985, her skin in the treatment area, her right hip, began to redden in a parallel striped pattern. The reddening did not immediately follow treatment with the Therac-25 because it generally takes at least several days before the skin reddens and/or swells from a radiation overexposure.

She continued Therac treatment until January 6, 1986 despite the reddening, since it was not determined that the reddening was an abnormal reaction. Hospital staff monitored the skin reaction and searched unsuccessfully for possible causes for the striped marks.

The hospital sent a letter to AECL and spoke on the phone with AECL's technical support supervisor, who later sent a written response stating, "After careful consideration, we are of the opinion that this damage could not have been produced by any malfunction of the Therac-25 or by any operator error." The hospital staff dismissed the skin/tissue problem as "cause unknown," partly due to the response from AECL, and partly because they knew AECL had already installed additional safety devices to their Therac-25 machine in September 1985.

Upon investigation in February 1987, the Yakima staff found the victim to have a chronic skin ulcer, dead tissue, and constant pain in her hip, providing further evidence

for a radiation overexposure. She underwent surgery and skin grafts, and overcame the incident with minor disability and some scarring related to the overdose.

March 22, 1986: Voyne Ray Cox. In March of 1986, AECL was notified that the Therac-25 unit in Tyler, Texas had overdosed a patient. They sent both a local Texas engineer and an engineer from their Canada home office to investigate the incident the day after it occurred. They spent a day running tests on the machine but could not reproduce the specific error. The AECL engineer suggested that perhaps an electrical problem had caused the accident. This was in part based on the patient's report that the accident produced a "frying" sound from the machine and significant heat. The engineer also said that AECL knew of no accidents involving radiation overexposure with the Therac-25.

At the East Texas Cancer Center (ETCC) in Tyler, Texas, 33-year old Cox was to receive his ninth Therac-25 radiation therapy session after a tumor had been successfully removed from his left shoulder. By this time the Therac 25 had been in successful operation at Tyler for two years, and 500 patients had been treated with it.

The Therac-25 operator left the radiation room to begin the treatment as usual. As she was typing in values, she made a mistake and used the "cursor up" key to correct it. Once the values were correct, she hit the "B" key to begin treatment, but the Therac-25 machine shut down after a moment, and the message "Malfunction 54" showed on the control room monitor. The machine indicated that only 6 of the prescribed 202 units of radiation had been delivered. The screen of the console showed that this shut down was a "treatment pause" which indicated a problem of low priority (since little radiation had been delivered). The operator hit the "P" key to proceed with the therapy, but after a moment of activity, "Malfunction 54" appeared on the Therac control screen again.

The operator was isolated from Cox because the Therac-25 operates from within a shielded room. On this day at the ETCC, the video monitor was unplugged and the audio monitor was broken, leaving no way for the operator to know what was happening inside. Cox had been lying on the treatment table, waiting for the usually uneventful radiation therapy, when he saw a bright flash of light, heard a frying, buzzing sound, and felt a thump and heat like an electric shock.

Cox, knowing from his previous 8 sessions that this was not normal, began to get up from the treatment table when the second "attempt" at treatment occurred. This time the electric-like jolt hit him in the neck and shoulder. He rolled off the table and pounded on the treatment room door until the surprised Therac-25 operator opened it. Cox was immediately examined by a physician, who observed reddening of the skin but suspected only an electric shock. Cox was discharged and told to return if he suffered any further complications.

The Fritz Hager, the hospital physicist was called in to examine the Therac-25, but no problems were found. The Therac-25 was shut down for testing the next day, and two AECL engineers, one from Texas and the chief engineer, Don Knott, from the home office in Canada, spent a day at the ETCC running tests on the machine but could not reproduce a Malfunction 54. Don Knott explained that the Therac-25 was unable to overdose a patient and also said that AECL had no knowledge of any overexposure accidents by Therac-25 machines. An independent engineering firm checked out the

electric shock theory and found that the machine did not seem capable of delivering an electric shock to a patient. The Therac-25 was put back into use on April 7, 1986.

Local management at the Cancer Center consulted with their superiors about how to respond. Because of legal ramifications, both felt it best not to notify others of the incident. Fritz Hager, however, convinced then that it was wise to notify the Texas Radiation Control Board, who also then notified the FDA.

Cox's condition worsened as he lost the use of his left arm and had constant pain and periodic nausea and vomiting spells. He was later hospitalized for several major radiation-induced symptoms (including vocal cord paralysis, paralysis of his left arm and both legs, and a lesion on his left lung). Cox died in August of 1986 due to complications from the radiation overdose.

April 11, 1986: Verdon Kidd. On April 11th of 1986, about 2 weeks after Cox's overdose/electric shock, Verdon Kidd was being treated at the Tyler, Texas facility for skin cancer on the side of his face. The same Therac operator who had treated Cox was treating Kidd. As the operator prepared to administer the Therac treatment from the control room, she used the "cursor up" key to correct an error in the treatment settings. She then began treatment using the "B" key.

The Therac-25 shut down within a few seconds, making a noise audible through the newly repaired intercom. The Therac monitor read "Malfunction 54." The operator rushed into the treatment room and found Kidd moaning for help. He said that his face was on fire. Fritz Hager, the hospital physicist was called. Kidd said that something had hit the side of his face, and that he had seen a flash of light and heard a sizzling sound.

After this second accident at the hospital, Hager took the Therac-25 out of service and called AECL. He worked with an assistant, and later with the Therac operator who had been administering treatment to both Cox and Kidd when the accidents occurred. Hager and the operator were eventually able to reproduce a Malfunction 54. They found that the malfunction occurred only if the Therac-25 operator rapidly corrected a mistake. The memo *How to produce a malfunction 54* from the medical physicist is available in the resources section.

Hager notified AECL of this discovery and AECL was eventually able to reproduce the error. AECL advised all Therac-25 users to physically remove the up-arrow key as a short-term solution. In their notification, they did not specifically mention the overdoses, but referred generally to danger that might occur from the editing process.

Fritz Hager was not satisfied with this response, and telephoned personally all the sites in the US and Canada that were using Therac-25. One by one, reports trickled back the next day that other medical physicists were able to reproduce the problem, confirming that it was a design flaw in Therac-25.

AECL filed a medical device report with the FDA on April 15, 1986 to notify them of the circumstances that produced the two Tyler accidents. At this point, the FDA, having been notified of the first Tyler accident by the hospital, declared Therac-25 defective and ordered the firm to contact all sites that used the machine, investigate the problem, and submit a report called a corrective action plan. AECL contacted all sites and recommended a temporary fix involving removing some keys from the keyboard at the computer console and notifying them in general terms of the error.

The FDA was not satisfied with the notification that AECL gave sites, and in May 1986 required AECL to re-notify all sites with more specific information about the defect in the product and the hazards associated with it. AECL was also at this time involved in meetings with a "user's group" of Therac-25 sites to help formulate its corrective action plan. After several exchanges of information among AECL and the FDA (in July, September, October, November, and December of 1986), AECL submitted a revised corrective action plan to the FDA.

The FDA worked in conjunction with AECL to identify the software problem and correct it. The FDA also requested that AECL change the machine in several other ways to clarify the meaning of malfunction error messages and to shut down treatment after any single large radiation pulse or interrupted treatment so that multiple overdoses were less likely.

Over the next three weeks Verdon Kidd became very disoriented and then fell into a coma. He had a fever as high as 104 degrees and had suffered neurological damage. He died on May 1, 1986.

January 17, 1987: Glen Dodd. Glen Dodd was at the Yakima Valley Memorial Hospital on January 17, 1987 to receive three sets of radiation treatment from the Therac-25.

The first two treatments went as planned. Dodd received 7 rads (radiation absorbed dose), 4 rads followed by 3 rads of radiation to take pictures of internal structure. The Therac-25 operator then entered the room and used the Therac-25's hand control to verify proper beam alignment on Dodd's body. Dodd's final dose of the day was to be a moderate 79-rad photon treatment.

The operator pressed a button to command the Therac to move its turntable to the proper position for treatment. Outside the treatment room, the Therac-25's control console read "beam ready," and the operator pressed the "B" key to turn the beam on. The beam activated, but the Therac-25 shut down after about 5 seconds. The console indicated that no dose had been given, so the operator pressed "P" to proceed with the treatment.

The Therac-25 shut down again, listing "flatness" as the reason for treatment pause. Dodd said something over the intercom, but the operator couldn't understand him. The operator went into the treatment room to speak with Dodd. He told the operator that he had felt a "burning sensation" in the chest. The operator's console displayed only the total dose of the two earlier treatments (7 rads).

Later that day, Dodd developed a skin burn over the treatment area. Four days later the burn was striped in a manner similar to that of the anonymous victim's burn after she had been treated at Yakima the year before.

AECL investigated the accident. All users were again told to visually confirm turntable setting before proceeding with any treatment. Given the information, it was suspected that the electron beam had come on when the turntable was in the field light position. AECL could not reproduce the error.

Later that week, AECL sent an engineer to Yakima to investigate. The hospital physicist had also been running tests. They eventually discovered a software flaw and fixed it. AECL engineers estimated that Dodd received between 8,000 and 10,000 rads instead of the prescribed 86. In February, 1987, the FDA and its Canadian counterpart

cooperated to require all units of Therac-25 to be shut down until effective and permanent modifications were made.

Glen Dodd died in April 1987. He had been suffering from a terminal form of cancer before the Therac accident, but it was determined that his death was primarily caused by complications related to the radiation overdose, not the cancer.

After another 6 months of negotiation with the FDA, AECL received approval for its final corrective action plan. This plan included numerous software fixes, the installation of independent, mechanical safety interlocks, and a variety of other safety related changes.

Government and FDA response to the Accidents

The Therac-25 case pointed to significant weak links in communication between the FDA, medical device manufacturers, and their customers or users. Users were not required to report injuries to any government office, or to the manufacturers of the devices that had caused injury.

The Food and Drug Administration (FDA) was created when Congress passed the Food and Drugs Act in 1906. This act was the first of a series of laws and amendments that gave the FDA jurisdiction over the regulation of foods and patent medicines. In 1938, Congress strengthened and expanded the FDA, to include the regulation of therapeutic and medical devices within its jurisdiction.

The FDA's Bureau of Medical Devices and Diagnostic Products was created in 1974, and soon operated in conjunction with the Medical Devices Amendments of 1976. The amendments helped to clarify the logistics of the regulation of medical devices, and required the FDA to "ensure their safety and effectiveness."

Radiation had been recognized as a health hazard since before World War I, and the FDA monitored the health risks that radiation emitting products posed to America's workers and consumers. As FDA's responsibilities for monitoring radiological devices grew, a bureau within the FDA called the Center for Devices and Radiological Health (CDRH) was established.

In 1980 the FDA's budget had swelled to over $320 million, with a staff of over 7,000. Many bureaus controlled areas such as biological drugs, consumer products, public health standards, and veterinary medicines.

FDA approved medical devices before they "went to market." This was called Pre-Market Approval and was a somewhat complex process. In the FDA Pre-market Approval scheme, devices were organized into three classes, as established by the 1976 Medical Device Amendments.

- *Class I devices*, These “present minimal potential for harm to the user” and are subject to "general controls” such a listing the device with the FDA. Examples include elastic bandages, examination gloves, and hand-held surgical instruments.
- *Class II devices*, such as syringes and hearing aids are those that "require performance standards in addition to general controls" to ensure safety and effectivenss.
- *Class III devices* like heart valves and pacemakers are devices that “support or sustain human life, are of substantial importance in preventing impairment of human health, or which present a potential, unreasonable risk of illness or injury”

These devices are required to undergo pre-market approval as well as complying with general controls

FDA approved Class III devices for market in one of two ways:

1. Proof of Pre-market Equivalence to another device on the market, termed 501(k)
2. OR Pre-market Approval (Rigorous Testing)

As a kind of grandfather provision, if a company could show Pre-market Equivalence (proof that a new product was equivalent to one already on the market before the 1976 amendments took effect), the new product could be approved by FDA without extensive, costly, rigorous testing. In 1984 about 94% of medical devices came to market through Pre-market Equivalence.

If a product was not equivalent to one that was already on the market, FDA required that the product go through extensive testing to gain Pre-market Approval. In 1984 only about 6% of medical devices were required to go through this testing.

Thus, it was clearly in the interest of medical device producers to show that their product had pre-market equivalence. Since Canadian Medical Company (CMC), designed the Therac-25 software based on software used in the earlier Therac-20 and Therac-6 models, Therac-25 was approved by FDA under Pre-market Equivalency.

A 1983 General Accounting Office (GAO) report criticized the FDA's "adverse experience warning system" as inadequate. FDA had published reports about potential hazards, including reports in their own newsletter, *The FDA Consumer*. The FDA implemented the mandatory medical-device reporting rule after Congress passed the Medical Device Reporting Legislation in 1984. This rule required manufacturers to report injuries and problems that could cause injuries or death.

Before 1986, users of medical devices (hospitals, doctors, independent facilities) were not required to report problems with medical devices. Instead, under the medical device reporting rule, manufacturers of these devices were required to report problems. The idea was that manufacturers would be the first to hear about any problems with the devices they made and that therefore reports would be timely. In addition, manufacturers would be most likely to have the correct information needed about a device to help resolve difficulties.

In the mid-1980s, the FDA's main enforcement tools for medical devices already on the market were publicity. The FDA could not force a recall; it could only recommend one. The CDRH (Center for Devices and Radiological Health monitors radiological devices) issues its public warnings and advisories in the Radiological Health Bulletin. Before issuing a public warning or advisory, the FDA could negotiate with manufacturers in private (and in the case of Therac 25, with regulatory agencies in Canada). In response to reports of problems with a medical device, the FDA could, in increasing order of severity:

- Ask for information from a manufacturer.
- Require a report from the manufacturer.
- Declare a product defective and require a corrective action plan (CAP).
- Publicly recommend that routine use of the system on patients be discontinued.
- Publicly recommend a recall.

In deciding on the response to a problem with a device, FDA needed to consider:

- Safety of the public.
- Safety of users of the device.
- Need for medical treatment with the device.
- Impact of the decision on the individual manufacturer.
- Impact of the decision on the medical device industry.

A 1986 GAO study found 99% of injuries caused by medical devices were not reported to the FDA. At that time, hospitals reported only about 51% of problems to the manufacturer. The hospitals mostly reported dealing with problems themselves. Problems were mainly the result of wear and tear on machines and design flaws.

The breakdown in communication with hospitals and clinics using medical devices prevented FDA from knowing about the isolated and recurring problems with the Therac-25 until after two deaths occurred in Tyler, TX.

Even when the FDA became aware of the problem, they did not have the power to recall Therac-25, only to recommend a recall. After the Therac-25 deaths occurred, the FDA issued an article in the Radiological Health Bulletin (Dec. 1986) explaining the mechanical failures of Therac-25 and explaining that "FDA had now declared the Therac-25 defective, and must approve the company's corrective action program."

After another Therac-25 overdose occurred in Washington state, the FDA took stronger action by "recommending that routine use of the system on patients be discontinued until a corrective plan had been approved and implemented" (Radiological Health Bulletin, March 1987). AECL was expected to notify Therac-25 users of the problem, and of FDA's recommendations.

After the Therac-25 deaths, the FDA made a number of adjustments to its policies in an attempt to address the breakdowns in communication and product approval. In 1990, health- care facilities were required by law to report incidents to both the manufacturer and FDA.

AECL Medical Goes Independent

AECL Medical, the division of AECL that designed and manufactured Therac-25 has become an independent private Canadian company, Theratronics. Theratronics was subsequently purchased by MDS Nordion, another Canadian manufacturer of medical devices. They still make radiation therapy machines.

Time Line

1970s	
Early1970's	AECL and a French Company (CGR) collaborate to build Medical Linear Accelerators (linacs). They develop Therac-6, and Therac-20. (AECL and CGR end their working relationship in 1981.)
1976	AECL developes the revolutionary "double pass" accelerator which leads to the development of Therac-25.
1983	
March, 1983	AECL performs a safety analysis of Therac-25 which apparently excludes an analysis of software.
July 29,1983	In a PR Newswire the Canadian Consulate General announces the introduction of the new "Therac 25" Machine manufactured by AECL Medical, a division of Atomic Energy of Canada Limited.
1984	
ca. Dec. 1984	Marietta Georgia, Kennestone Regional Oncology Center implements the new Therac-25 machine.
1985	
June 3, 1985	Marietta Georgia, Kennestone Regional Oncology Center Katherine (Katy) Yarbrough, a 61-year-old woman is overdosed during a follow-up radiation treatment after removal of a malignant breast tumor. Tim Still, Kennestone Physicist calls AECL asking if overdose is possible; three days later he is informed it is not.
July 26, 1985	Hamilton, Ontario, Canada. Frances Hill, a 40-year-old patient is overdosed during treatment for cervical carcinoma. AECL is informed of the injury and sends a service engineer to investigate.
November 3, 1985	Hamilton Ontario patient dies of cancer, but it is noted on her autopsy that had she not died, a full hip replacement would have been necessary as a result of the radiation overdose.
November 8, 1985	Letter from CRPB to AECL requesting additional hardware interlocks and changes in software. Letter also requested treatment terminated in the event of a malfunction with no option to proceed with single key-stroke. (under Canada's Radiation Emitting Devices Act.)
November 18, 1985	Katy Yarbrough files suit against AECL and Kennestone Regional Oncology Center. AECL informed officially of Lawsuit.
December 1985	Yakima Valley Memorial Hospital, Yakima Washington. A woman being treated with Therac-25 develops erythema on her hip after one of the treatments.

1986	
January 31, 1986	Staff at Yakima sends letter to AECL and speak on the phone with AECL technical support supervisor.
February 24, 1986	AECL technical support supervisor sends a written response to Yakima claiming that Therac-25 could not have been responsible for the injuries to the female patient.
March 21, 1986	East Texas Cancer Center, Tyler Texas. Voyne Ray Cox is overdosed during treatment on his back. Fritz Hager notifies AECL. Company suggests some tests and suggests hospital might have an electrical problem. AECL claims again that overdoes is impossible and that no other accidents have occurred previously.
March 22, 1986	Ray Cox checks into an emergency room with severe radiation sickness. Fritz Hager calls AECL again and arranges for Randy Rhodes and Dave Nott to test Therac. They travel to Texas and test Therac but find nothing wrong.
April 7, 1986	ETCC has investigated electrical problem possibility, finding none, put Therac-25 back in service.
April 11, 1986	East Texas Cancer Center. Another Verdon Kidd is overdosed during treatments to his face. Operator is able to explain how Malfuction 54 was achieved. Fritz Hager tests computer's readout of no dose, and discovers the extent of the overdoses. Hager spends weekend on phone with AECL explaining findings.
April 14, 1986	AECL files report with FDA. AECL sends letter to Therac-25 users with suggestions for avoiding future accidents, including the removal of the up-arrow editing key and the covering of the contact with electrical tape.
May 1, 1986	Verdon Kidd, who was to have received treatments to left ear dies as a result of acute radiation injury to the right temporal lobe of the brain and brain stem. He is the first person to die from therapeutic radiation accident.
May 2, 1986	FDA declares Therac-25 defective, and their "fix" letter to users inadequate. FDA demands a CAP from AECL.
June 13, 1986	AECL produces first CAP for FDA.
July 23, 1986	FDA has received CAP, asks for more information.
August, 1986	Therac-25 users create a user group and meet at the annual conference of the American Association of Physicists in Medicine
August, 1986	Ray Cox, overdosed during back treatment, dies as a result of radiation burns.
September 23, 1986	Debbie Cox and Cox family file lawsuit
September 26, 1986	AECL provides FDA with more information.

October 30, 1986	FDA requests more information.
November 1986	Physicists and engineers from FDA's CDRH conducted a technical assessment of the Therac-25 at AECL manufacturing plant in Canada (R.C. Thompson).
November 12, 1986	AECL submits revision of CAP.
December	Therac-20 users notified of a software bug.
December 11, 1986	FDA requests more changes to CAP.
December 22, 1986	AECL submits second revision of CAP.
1987	
January 17, 1987	Second patient, Glen A. Dodd, a 65-year-old man, is overdosed at Yakima.
January 19, 1987	AECL issues hazard notification to all Therac-25 users and told them to visually confirm the position of the turntable before turning on beam.
January 26, 1987	Conference call between AECL quality assurance manager and Ed Miller of FDA. AECL sends FDA revised test plan. AECL calls Therac users with instructions on how to avoid beam on when turntable is in field light position.
February 3, 1987	AECL announces additional changes to Therac-25
February 6, 1987	Ed Miller calls Pavel Dvorak of Canada's Health and Welfare department with news that FDA will recommend that all Therac 25 units be taken out of service until CAP is completed.
February 10, 1987	FDA sends notice to AECL advising that Therac is defective under US law and requesting AECL to notify customers that it should not be used for routine therapy. Canadian Health Protection Branch does the same.
March 1987	Second User Group Meeting
March 5, 1987	AECL sends third revision of CAP to FDA.
April 1987	Glen A. Dodd, overdosed at Yakima, dies of complications from radiation burns to his chest.
April 9, 1987	FDA asks for additional information regarding third CAP revision.
April 13, 1987	AECL sends update of CAP and list of nine items requested by users at March meeting.
May 1, 1987	AECL sends fourth revision of CAP to FDA as a result of FDA commentary at user meeting.
May 26, 1987	FDA approves fourth CAP subject to final testing and analysis.
June 5, 1987	AECL sends final test plans to FDA along with safety analysis.

July, 1987	Third Therac-25 User Group Meeting
July 21, 1987	AECL sends final (fifth) CAP revision to FDA.
1988	
January 28, 1988	Interim safety analysis report issued from AECL.
November 3, 1988	Final safety analysis report issued.

Perspective Pieces.

Atomic Energy Canada Limited and its Corrective Action Plans

The Tyler, TX bug fix.
AECL was notified by ETCC medical physicist Frtiz Hager that the error that produced both Tyler, TX incidents was now reproduceable. Don Knott, the chief engineer at AECL, after some effort was also able to reproduce it on the machine they had there (see *historical narrative* for more detail). AECL management then filed a medical device report with the FDA, notifying them of the problem. After some negotiation with the FDA, AECL managers agreed to:

1) Notify all sites about the specific hazards associated with this defect
2) Recommend a temporary fix for the defect until a permanent fix had been implemented
3) Begin conversations with a Therac-25 "user group" to plan for implementation of the fixes
4) Implement and distribute a permanent fix within a set period of time

The temporary fix involved removing the up-arrow key from the keyboard so operators could not edit the parameters in a way that would cause the defect to appear (see historical (see "an advanced technical lesson" for information on precisely how the defect occurred). AECL's permanent fix involved changes to the software that:

1) Removed the specific error that cause the race condition and defect
2) Gave more clear feedback in the interface regarding the meaning of error messages
3) Included an automatic shutdown after any single large pulse of radiation, so that multiple large doses were less likely.

Decision Point 1: Is it fixed (See exercises)
Decision Point 2: Constructing a Corrective action plan (See exercises)

Fritz Hager and local action in Tyler, TX

What Facilities are like

Cancer treatment facilities are often housed in large hospitals, but some are stand-alone cancer treatment centers, like the East Texas Cancer Center where Fritz Hager worked. ETCC was a part of a larger ono-profit organization that ran the cancer treatment center, a hospital, and several other medical facilities..

During the time of Therac-25 (the mid 80s) a well-equipped treatment facility might have 3 different machines. The machines would be capable of producing different kinds of radiation, different strengths of beam, and capable of different kinds of exposure to the patient. Each of these machines would cost, for the machine alone, between 1 and 2 million dollars. In addition, special housing for each machine is needed, with shielding in the walls, adequate power supply, video and intercom links, etc.

ETCC had, in fact, three radiation therapy machine: a Varian, a Therac-6, and a Therac-25. Both the Varian and the Therac-6 were fixed-mode machines that only delivered Y-Rays, while the Newer Therac-25 could treat with either X-rays or an electron beam (see Historical documents for more detail). Using al three machines, ETCC could treat about 100 patients a day.

Operators would be needed to run each machine. For larger facilities, as at ETCC, a supervisor of the operators, with more training and experience might be needed. In addition, at least one MD specialist in cancer radiation therapy (a Radiation Oncologist) would be required. Finally, a medical physicist would be needed to maintain and check the machines regularly. Some facilities contract out the services of a medical physicist, But at ETCC, Fritz Hager was in charge of the radiation therapy machine and several other machine at related sites that used radiation. Finally, all the support personnel for these specialists (nurses, secretaries, administrative staff, people to handle billing and paperwork, janitorial staff, etc.) are required.

Machine Support and Maintenance

Medical Linear Accelerators do age over time, and older machines often produce more errors. Five to ten years is a reasonable life span for a machine. Thus, simply to maintain a set of three medical linear accelerators, an institution can expect to spend 1 to 2 million dollars every third year.

Sometimes errors can be resolved and a machine kept longer using software upgrades or upgrades or retrofits of machine parts. The companies that sell linear accelerators charge maintenance contracts that can include different levels of support. Because of monetary constraints, sometimes facilities are forced to choose between software updates, manuals, and training for operators and physicists. All this is in addition to the price of the machine itself.

Production Pressures
Production pressures are always present when an expensive medical technology is being used. These very expensive machines need to treat enough patients to pay for themselves over their lifetime. Another kind of production pressure is generated because of concern for the patient. Patients' schedules require treatments on certain days and it disrupts the patients' lives and slows down their treatment to have to reschedule them for another day while the machine is being checked out.

These production pressures generate the desire to "push patients through." If a machine gives only a portion of the prescribed dose, an operator will often repeat the treatment with enough radiation to add up to the total prescribed dose. Of course, because of liability issues and concerns for patient welfare, this can only be done when it is thought safe.

One of the advantages of the significant computerization of the Therac 25 and Therac-6 machines was that setup for treatment could be done much more quickly. This allowed the operator more time to speak with the patient and interact with them about their health concerns. In addition, this increased efficiency allowed more patients to be scheduled during a day. Thus, more patients could be treated, but the atmosphere was not reduced to that of a factory.

Liability and Trust
Facilities that run medical linear accelerator are surely concerned about liability for injury to patients that might occur. Insurance, for medical providers, is quite expensive and errors in treatment can result in lawsuits, which in turn produce increases in insurance premiums. Standard practice in litigation is to "sue everyone with deep pockets." This means that even if an error is the result of poor design of a linear accelerator, the facility itself will be sued simply because they were involved: they have insurance and thus "deep pockets."

But it is in the interest of facilities to reduce errors without the threat of lawsuits. When a treatment must be restarted several times because of errors, it may reduce patient confidence in the facility. This can mean patients moving to another facility with which they are more comfortable. Finally, medical professionals are in their business because they want to help people and have the knowledge and skill to do so. So a primary motivation of medical professionals is patient welfare. But the pressure from lawsuits often works against the interest in reporting, identifying, and resolving medical errors.

For instance, after the first incident at ETCC, involving Voyne Ray Cox, Fritz Hager was unable to interview Cox to talk about incident or the injury. He had to rely on second hand reports about it from the operator and from the doctor who saw Cox immediately after the injury. Cox had been taken to another facility by his relatives, who were threatening lawsuits, and thus would not speak with Hager as he tried to identify the source of the problem. Without this information, Hager had to bring in engineers (both from AECL and an electrical firm) to check the machine. The second-hand reports suggested Cox had received a shock, and this lead the engineers, and Hager to pursue this

avenue. Don Knott, the AECL engineer, had told Hager that an overdose was impossible, so this theory was discarded.

Decision Point 3: Report the Problem or Deal with it Internally
Decision Point 4: Going Public on your own.

Supporting Documents.

Historical Documents.

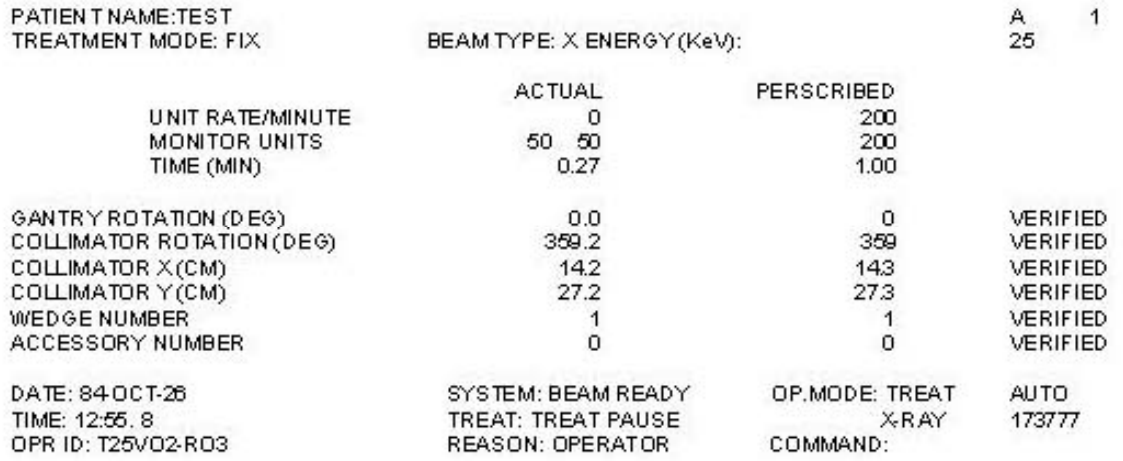
How to Produce a Malfunction 54 on a AECL Therac-25 Linear Accelerator

This statement was written by Fritz Hager, the East Texas Cancer Center physicist, after he discovered how to reproduce the Malfunction 54 error:

Enter the room and set up the machine for an electron beam treatment by selecting a field size and installing the trimmers. Press the set button. Leave the room and close the door. At the control console proceed to the patient set-up display. For Mode enter "X". The machine will default to 25 MeV and go to dose rate of 250 rads/min. Use return key to go to dose. Enter 200. Use return key to go to time. Enter 0.8 min. Use the return key to rapidly advance to the bottom of the display. Immediately use the up arrow to move from the bottom of the display. You are now in the edit mode. Use the up arrow to go to the top of the display and change the mode "X" to "E" for electrons. Change the energy from 25 to 10. Use the return key to go back down to the bottom of the display. Wait for the "beam ready" message then type "B" return. The unit will have no indications on dose rate or dose 1 or dose 2 for about 3 to 4 seconds. Then the dose rate will flash 550 to 575 for one cycle and return to zero. Dose 1 and Dose 2 will count to -6. A malfunction 54 message will appear at the bottom of the display. You have just delivered a dose of approximately 25,000 rads of 25 MeV electrons in less than two seconds.

Maps, Tables, & Figures.

A Screen Shot of the Interface for Therac-25



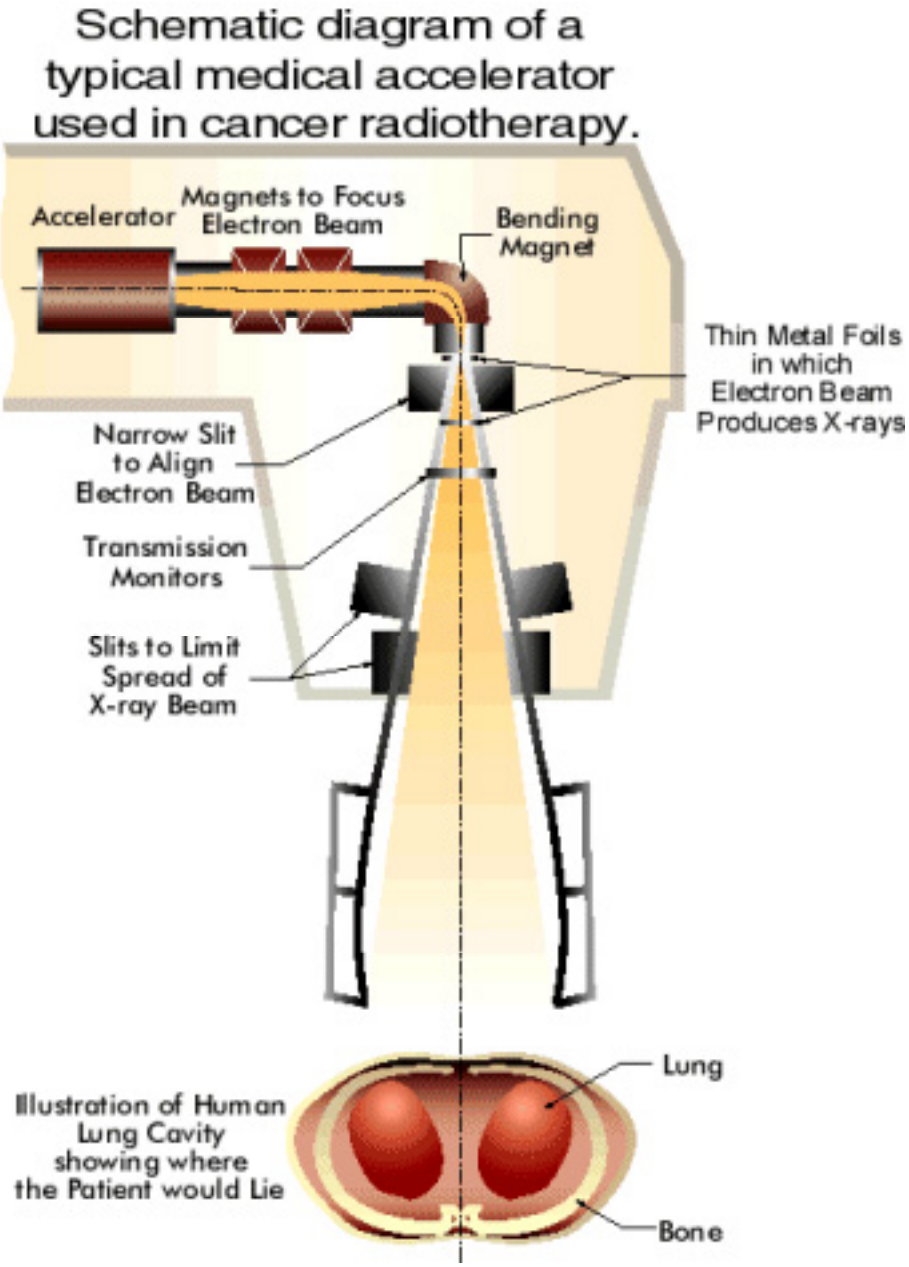
PATIENT NAME:TEST
TREATMENT MODE: FIX
BEAM TYPE: X ENERGY (KeV):
A 1
25

	ACTUAL	PERSCRIBED	
UNIT RATE/MINUTE	0	200	
MONITOR UNITS	50 50	200	
TIME (MIN)	0.27	1.00	
GANTRY ROTATION (DEG)	0.0	0	VERIFIED
COLLIMATOR ROTATION (DEG)	359.2	359	VERIFIED
COLLIMATOR X (CM)	14.2	14.3	VERIFIED
COLLIMATOR Y (CM)	27.2	27.3	VERIFIED
WEDGE NUMBER	1	1	VERIFIED
ACCESSORY NUMBER	0	0	VERIFIED

DATE: 84-OCT-26
TIME: 12:55. 8
OPR ID: T25V02-R03
SYSTEM: BEAM READY
TREAT: TREAT PAUSE
REASON: OPERATOR
OP.MODE: TREAT
X-RAY
COMMAND:
AUTO
173777

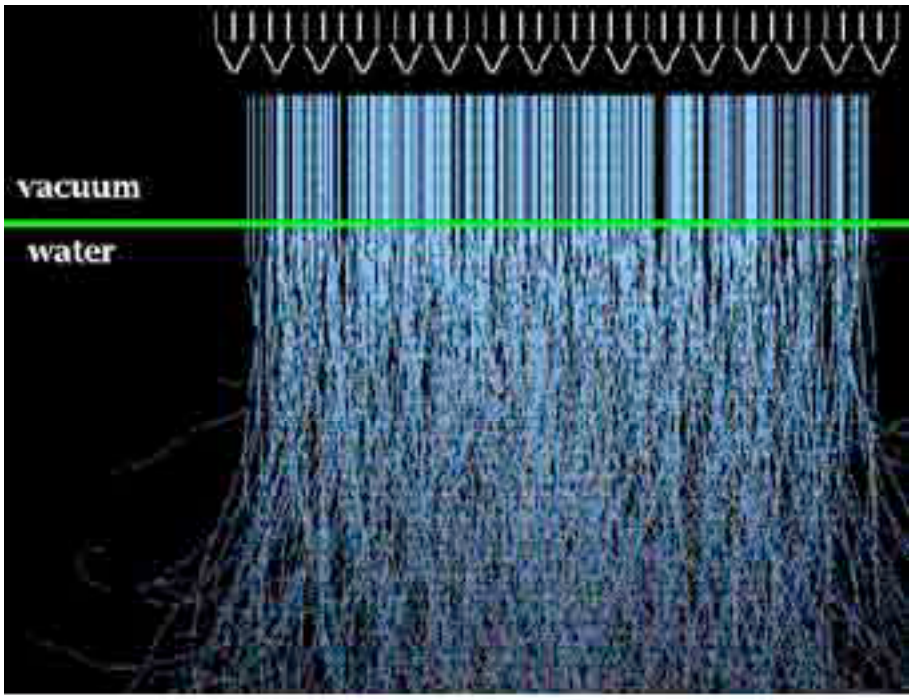
A schematic of a generic medical accelerator for radiation therapy.

This is not the specific therac-25 machine, but an example of the standard components in a medical linear accelerator.



A model of radiation dispersion in water.

This model allows you to think about how radiation from a medical linear accelerator might penetrate the body to treat a cancer. This is a model by the Stabford Linear Accelerator Center of dispersion of radiation from a beam through water.



SLAC has an excellenct model demontrating low density beam penetration.This photo is for demonstration only and has **no** involvement with Therac.(Photo courtesy of SLAC)

A model of the Therac-25 machine in a dedicated room.

You can clearly see here that the Therac-25 machine is part of a complex physical setup.

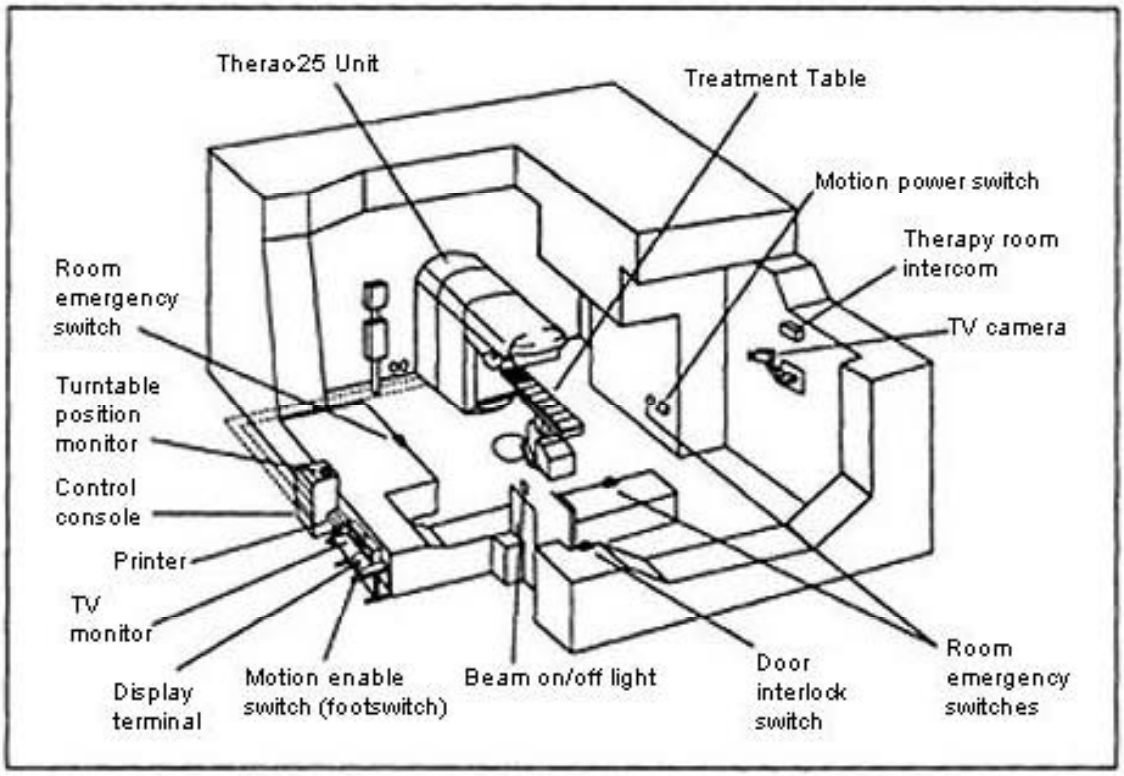


Figure 1. Typical Therac-25 facility

A Schematic of the turntable assembly

This placed different filters (or no filter) in front of the beam.

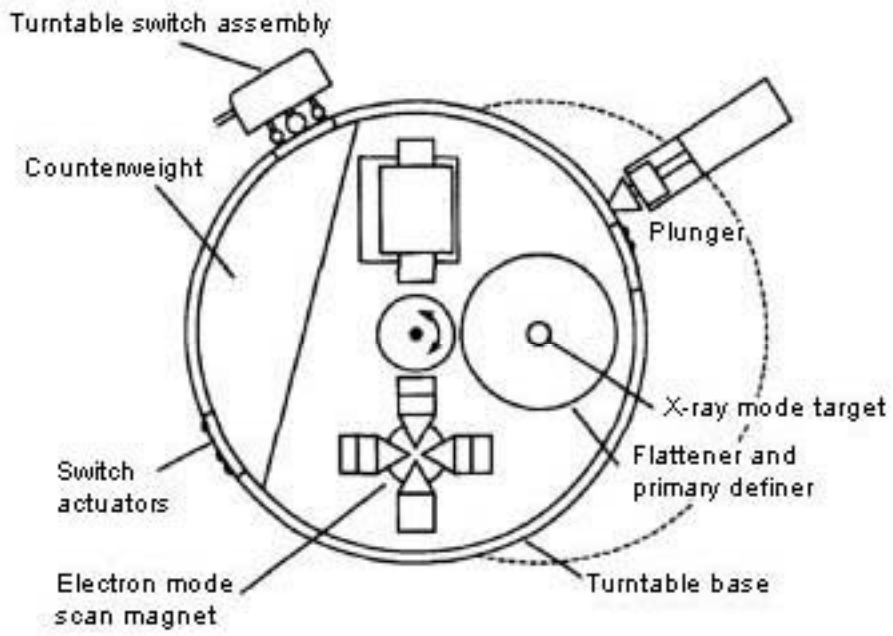


Figure B. Upper turntable assembly

Other resources.

A beginning technical lesson from the case

One of the software errors was documented as the cause of a death at the Yakima, WA treatment center. It is a simple enough error for students with little CS background to understand, though it raises profound ethical issues about the reuse of software. This error involved incrementing a shared variable called Class3 instead of doing a Boolean. Class3 indicated whether a particular piece of equipment that shaped and limited the beam (the collimator) was or was not in the right position. If the collimator's position was correct, Class3 was set to zero, otherwise it was set to some non-zero integer by the simple expedient of incrementing it. Class3, however, was stored in a single byte, and so every 256 times it was incremented, it would overflow and register zero. If the operator pressed the set key at precisely that moment, the machine would schedule a treatment without the shaping device in the correct place, causing a race condition which resulted in an overdose. This was a rare enough occurrence that it was difficult to detect.

The contemporary milieu in which Therac programmers worked informs our understanding of the Yakima incrementation error. Incrementing a variable to change its state from zero (FALSE) to non-zero (TRUE) was a common and standard practice in the day. Assembly language programmers particularly made use of such tricks to save a few precious CPU cycles in their code. The problem with incrementation to change a boolean variable arises when that variable overflows. In the Yakima problem, such overflows, together with bad timing (race conditions), had fatal consequences. Using a two-byte PDP-11 word instead of a single byte for that variable would have made these overflows 256 times less likely (one in 65,536 instead of one in 256); a four-byte longword might never have overflowed (one in over four billion). But longwords were not available in the "16-bit" PDP-11 architecture, unless one built a multi-precision integer variable oneself (at a cost to performance); and the first inclination of any assembly programmer would be to try to get away with a single byte, applying the same minimal-usage philosophy to memory space as to CPU execution time.

Furthermore, it is unclear whether the requirements for the early 70s (and more simple) Therac-6 software influence the decision to use incrementation instead of a boolean.. Was incrementation of an 8-bit variable to change its (boolean) state part of the reused Therac-6 software, perhaps intended for initial calibration only, but later used throughout the control of the more complicated dual mode Therac-25? We cannot tell without knowing more specifics about the Therac code and its evolutionary development. In any case, this hypothesis indicates the kind of imagination about possible effects when a bit of software is reused that we must expect from programmers. Such imagination is difficult enough to instill in present-day programmers, and was in very short supply among 1970s assembly language programmers.

The world has learned much about how to think about software reliability over the last three decades. It is now inconceivable that the FDA would approve a software-dependent device such as the Therac-25 with as little testing of the software functionality as that system received---in large part due to the experience gained from this very system and others that looked acceptable at the outset but later generated unanticipated problems. The word that little logical flaws such as incrementing to change boolean state could result in catastrophic failure was slow to trickle back to programmers.

The software reuse in this case are makes it clear that thoughtful documentation is important when software may be reused. Software designed for one situation (where there are hardware interlocks) may be reused in a different system. These shifting requirements can cause software to fail catastrophically as the context of its use changes. Software in one socio-technical context can operate flawlessly and can, in another context, produce significant harm. This is what "best practices" in software design processes are designed to handle. All of these points make it clear how closely coupled technical and ethical issues are.

An advanced technical lesson from the case

The other documented software error was the cause of two deaths in the Tyler, TX treatment facility. It involved concurrent access to a shared variable, that set up a race condition resulting in a mismatch of beam energy with the turntable position. In this case, it involved monitoring and verifying treatment data input from the operator. The variable indicated whether data entry was completed or not. Under certain conditions, a race condition would occur with this variable and if the operator set initial parameters and later edited them the later edits would show on the screen but would not be detected by the portion of the software that implements the machine settings. This could result in the wrong piece of the turntable being in place with the high energy beam turned on, producing a massive overdose.

In technical terms, these race conditions would have been resolved by indivisible set and test operations. Explaining what this means and why it is important is a topic that comes up on operating systems courses today, but that was not widely known and only poorly understood in the early 70s when some of the Therac-6 software was produced. You can see that technical issues from your courses can have profound ethical implications.

A closer look at the staged development of the Therac-6, Therac-20 and Therac-25 software given "best" computing practices in the 1970s raises uncertainties about the overall responsibility for race condition failures in that system. In order to understand these uncertainties, students first need a review of (or introduction to) the technical issues.

A race condition in a software system arises when the correct operation of that system depends on timing of execution. Race conditions can only occur in systems that involve multiple tasks (or processes) that carry out their instructions concurrently, or at the "same time," perhaps through a timesharing operating system. To illustrate the subtle nature of race conditions in a multitasking system, consider Dijkstra's classic "dining philosophers" problem (1965): Imagine several individuals sitting around a dinner table, each of whom alternates between thinking and eating for indeterminate amounts of time. In order to eat, one of these "philosophers" requires two utensils. In Dijkstra's analogy, two forks were required to eat spaghetti; others have suggested chopsticks and Chinese food. Unfortunately, each philosopher must share his/her utensils with his/her neighbors. Imagine a circular table with 8 plates around the edge and 8 single chopsticks, one to the right of each plate. To eat, each philosopher must pick up one chopstick from each side of his or her plate. This means that the philosophers on either side of the dining philosopher will have to wait to eat (in the analogy, they spend the time thinking), since they each have access to only one chopstick. In this hypothetical problem, the philosophers represent multiple concurrent tasks, and the chopsticks represent computing resources (variables, input/output devices, etc.) that those processes must share.

Correctly instructing (programming) each philosopher-task about how to proceed is a non-trivial problem, as shown later. For example, in an attempted solution algorithm for which each philosopher receives a boolean-valued shared variable that is TRUE when that philosopher may safely pick up both utensils and eat, inopportune timing may lead one philosopher A to yield eating rights to a second philosopher B between the time when B checks for eating rights and the time when B begins to think, awaiting a "wakeup" from A; if A sent B's "wakeup" signal before B has begun to sleep, then B

might enter a state of permanent inactivity while unknowingly holding the sole right to eat. Thus, the correct behavior of the system of philosopher algorithms depends on timing---a race condition. Naïve efforts to repair this approach by saving "wakeups," sharing different variables, etc., simply move the bad-timing condition to another location in the multitasking algorithm. Furthermore, in practice, race conditions are among the most difficult system bugs to find, since one cannot duplicate the precise timing conditions that cause the transient error, except by chance, unless one already knows what that transient error is. In the Therac-25 case, the Tyler race condition did not show up unless particularly well practiced operators made editing changes very quickly.

The only effective way to combat race conditions is to avoid them. In the example above, one can avoid the lost "wakeup" by insuring that philosopher B cannot be interrupted between the step of checking its shared variable and the step of entering its thinking phase, i.e., by causing the system to perform both steps in a single indivisible operation. In this way, the philosopher is assured that the unfortunate timing error of having the wakeup call occur right after checking for it but before entering the thinking stage (when it would cause some action) is made impossible.

Now that we know what a race condition is, we can see in the Therac-25 how it can kill people. Therac-25 appeared in 1983, and its programmers modified the software from the earlier Therac-6 system, programmed in the early-1970s. Given the era when software for these systems appeared (the "state of the technical art" in the socio-technical system section), it seems unlikely that machine language programmers in 1972 would have sufficient knowledge to know how to avoid race conditions with concurrent processes. This makes the technical error more understandable. But it brings up the larger issue, again, of documentation of code when it is reused and of the effect of shifting requirements and socio-technical systems. It also makes clear the importance of including a carefully designed maintenance phase in the software life-cycle; Part of the reason the errors were not caught is that there was no clear mechanism for them to filter back to the appropriate people at AECL (or the FDA). Again, what look like technical issues become ethical ones. Students who learn about race conditions from the Therac-25 case know about the importance of a professional's responsibility to be aware of best practices and the state of the art in one's area of expertise, and the dangers of operating outside of their area of expertise.

Leveson & Turner say that "focusing on particular software bugs is not the way to make a safe system …The basic mistakes here involved poor software-engineering practices and building a machine that relies on the software for safe operation."[1] This is not the lesson that is often drawn from the case, but presenting the case in this fashion makes it clear technical decisions made by the programmers in the context of their historical environment had profoundly ethical implications for the eventual users of the system. In this way putting the case in historical perspective makes it clear that ethical decision making is inseparable from good software design methodology, and that *good* in this context deserves its double entendre.

Basics of Concurrency and Race Conditions

•Multitasking

- A *process* (or ***task***) is an execution of a program. Note that in a multiuser system, we would expect multiple processes executing various programs; it may be that

[1] Leveson and Turner, 38

two or more processes even execute the same program "simultaneously" through time sharing.

- ***Multitasking*** means using multiple tasks in a single system.
  Example: Dining philosopher's problem (see below), which involves multiple tasks sharing various resources.
- The use of multitasking leads to new and complicated kinds of computer bugs.
  Example: ***deadlock*** -- existence of a set of processes, each of which is ***blocked*** (unable to run) waiting for an event that can only be caused by another process in that set. See first attempted solution of dining philosopher's problem below.
- Multiple processes may require ***shared variables*** in order to carry out their work.
  Example: Two processes may use a shared "buffer" data structure for holding items that have been produced by one process but not yet consumed by another.

Example: Dining Philosophers Problem (E. Dijkstra, 1965)

- *N* processes share *N* resources in an alternating circular arrangement. Each process has two states (computing, interacting with resources); each needs exclusive access to its two resources while interacting with them.
- Dijkstra's statement: *N* philosophers sit down together to eat spaghetti. Each philosopher spends his/her time alternately thinking then eating. In order to eat, a philosopher needs two forks; each fork is shared with exactly one other of the philosophers. What procedure will allow all the philosophers to continue to think and eat?
- Algorithm 1. Each philosopher does the following:

```
repeat forever
   think
   pick up left fork (as soon as it's available)
   pick up right fork (as soon as it's available)
   eat
   return left fork
   return right fork
```

Issue: *Deadlock* occurs if each philosopher picks up his/her left fork simultaneously, because no philosopher then can obtain a right fork.

- Algorithm 2. Each philosopher does the following:

```
repeat forever
   think
   repeat
      pick up left fork (as soon as it's available)
      if right fork is available
         pick up right fork
      else
         return left fork
   until both forks are possessed
   eat
   return left fork
   return right fork
```

Issue: Although deadlock has been prevented, s*tarvation* occurs if each philosopher picks up his/her left fork simultaneously, observes that the right fork

is unavailable, then simultaneously returns the left fork and tries again, ad infinitum.

Race conditions

- A ***race condition*** exists in a system if the correct operation of that system's algorithm depends upon timing.
  Example: Third attempted solution of dining philosopher's problem below. This algorithm uses a shared array myturn[] in an effort to prevent problems like deadlock. But "bad luck" in terms of timing could lead to a deadlock after all.
- Example of an algorithm with race conditions: Algorithm 3 for Dining Philosophers. The philosophers use a shared variable myturn, an array of length *N*; if myturn[*p*] is *true* then it is (hopefully) safe for philosopher *p* to eat. (We use an array rather than a single variable location to allow for *concurrency*---multiple processes executing at the same time.)
  We assume that there are functions sleep() for causing a philosopher to doze off (blocked process) and wakeup(*p*) for waking up a philosopher *p*.
  Each philosopher does the following:

```
if (p == 1)
   myturn[p] = true
else
   myturn[p] = false
next = (p+1) % N
prev = (p+N-1) % N
repeat forever
   think
   if (!myturn[p])
      sleep()    /* to be awakened by prev philosopher */
   pick up left fork
   pick up right fork
   eat
   return left fork
   return right fork
   myturn[p] = false
   myturn[next] = true
   wakeup(next)
```

Issues: Less than maximal concurrency; race conditions

- The following illustration shows how bad timing might occur with the algorithm above.
-

```
philosopher p=1                philosopher p=2
                                +>
                                |       ...
                                |       // myturn[p] IS FALSE (p=2)
                                |       think
                                |       if (!myturn[p])
                               <+
   ...                          |
   myturn[p] = false //p=1      |
   myturn[next] = true          |
   wakeup(next)  // WAKEUP p=2  |
repeat forever                  |
   think                        |
```

```
   ...                          |
   if (!myturn[p])              |
      sleep()   // BLOCK!       |
                               +>
                               |         sleep() // BLOCK!
```

•

•

•The next illustration shows a second race condition during initialization.

```
philosopher p=1                philosopher p=2
                                <+
if (p == 1)                      |
   myturn[p] = true              |
else                             |
   myturn[p] = false             |
next = (p+1) % N                 |
prev = (p+N-1) % N               |
repeat forever                   |
   think                         |
   if (!myturn[p]) // FALSE      |
      sleep()                    |
   pick up left fork             |
   ...                           |
   myturn[p] = false             |
   myturn[next] = true           |
                                +>
                                |   if (p == 1)
                                |     myturn[p] = true
                                |   else
                                |     myturn[p] = false
                                |   // myturn[p] IS NOW FALSE
                                |   next = (p+1) % N
                                |   prev = (p+N-1) % N
                               <+
   wakeup(next)  // WAKEUP p=2 |
repeat forever                   |
   think                         |
   ...                           |
   if (!myturn[p])               |
      sleep()   // BLOCK!        |
                                +>
                                |   repeat forever
                                |      think
                                |      if (!myturn[p])
                                |         sleep() // BLOCK!
```

•Race conditions are *extremely* difficult to debug , since one must look for bad combinations of instruction execution order for multiple independent programs. The worst part about debugging race conditions is that the bugs are *intermittent* --- not usually repeatable on a subsequent run, as hard as you may try to replicate the running conditions.

- The best approach is to *avoid race conditions in the first place*: A race condition can only arise when there can be a "gap" between *retrieving* a shared variable's value and *using* that retrieved value. See dining philosopher race condition examples above. *Comment: persons who haven't had a course like Operating Systems are likely to be totally unaware of this race condition issue!*

- **Strategies for correct IPC**
  - Race conditions are problems in *interprocess communication (IPC)* or ***synchronization***, i.e., mechanisms and practices for creating correct systems having multiple processes.
  - To prevent race conditions, one needs some kind of *atomic* or ***indivisible*** operations that leave no possibility for timing "gaps."
  - An Operating Systems course examines several (equivalent) higher-level software strategies for correct IPC, including *semaphores*, *monitors* (cf. Java's synchronized objects), and *message passing*.
    Comment: Therac25 assembly programmers would not have any of these higher level solutions available unless they built them themselves --- an unlikely scenario for time-pressured programmers with only low-level programming resources who might not even have awareness of the issue, especially given the care that correct programming of one of these strategies would require.
  - There are also hardware solutions, such as having a *test and set lock (TSL)* instruction in the ISA (machine language) of the machine being used. In a TSL instruction, a memory value can be retrieved ("tested") and a new "lock" value substituted in a single indivisible machine instruction, preventing a "gap" between "testing" and "locking." Thoughtful use of a TSL instruction can correctly solve IPC problems.
  - Leveson and Turner's analysis: *"It is clear from the AECL documentation on the modifications that the software allows concurrent access to shared memory, that there is no real synchronization aside from data stored in shared variables, and that the "test" and "set" for such variables are not indivisible operations. Race conditions resulting from this implementation of multitasking played an important part in the accidents."*
  - *Comment: The Therac25 system used a standard, highly regarded computer produced by DEC (Digital Equipment Corporation) called the PDP11. The PDP-11 instruction set includes no TSL instruction. (There is also no SWAP instruction to interchange values of two independent memory locations, which could serve in place of a TSL instruction.) Thus, Therac25 programmers would have had to devise something else for correct synchronization.*

Interview with a Therac-4 Operator

This is a summary based on an interview we conducted with a Registered Therapy Technologist who has extensive experience operating medical linear accelerators. This individual currently manages a Radiation Therapy Department at a major United States hospital, and trains technicians to operate radiation therapy machinery. For privacy purposes, the true identity of this person will remain anonymous, and for the remainder of the article, we will refer to our interviewee as "Susan."

Susan operated a Therac-4 linear accelerator machine in the mid 1980's. At the time, Susan had recently graduated and was working at a University where the radiation therapy technology was fairly advanced. She enjoyed operating AECL's Therac machine because it was one of the first computerized linear accelerators. Looking back, Susan remembered that while operating the machines, she did not think much about whether there could be computer software "bugs" in the system. The technology was new, and she remembered trusting the machine's components and its designers.

When recalling the advantages of the new computerized machine, Susan reported being able to move more patients through during the day. She also remembered feeling good about the extra time she had to talk with patients when she was working with a computerized machine.

Susan learned about the Therac-25 incidents while attending a national radiation therapy conference in 1990. A radiation therapist who was also a lawyer gave a lecture on the Therac-25 accidents. He handed out newspaper articles about the incidents and spoke about how many times the therapists involved in the accidents attempted to resume treatment in spite of the error messages they received from the computer. The lecture focused on the question of how many attempts to resume treatment is too many? The lecturer and the participants discussed the possibility of establishing institutional policies and limits on the number of times an operator could resume treatment after having received an error message, such as the cryptic "malfunction 54" messages that the operator received during the two fatal accidents in Texas.

The problem, Susan reported, is that back in 1990, and today, there are no industry-wide standards or rules for these types of situations. Susan felt that she had been lucky to have always worked where there was a physicist available to provide help with the many error messages operators received. She also felt that in other clinics, where this kind of assistance is not available, there was, and still is, a great deal more pressure on therapists to just keep going despite the error messages. An operator might attempt, for example, to deliver the prescribed dose in 12 increments instead of 1 by continually clearing the faults generated by the computer. Susan stated that this type of activity happens all the time in medical radiation therapy, particularly in clinics where there is more pressure from the administration to keep patients moving through quickly.

Although Susan had been working with a AECL Therac machine at the time of the accidents, she did not remember receiving warning notices from AECL about the Therac-

related accidents. Susan believes that this is one aspect of the industry that has changed, possibly, in part due the Therac-25 accidents. At the present time Susan receives notices from the manufacturers of the linear accelerators used at her hospital whenever there is a linear accelerator malfunction, or even if there is a malfunction that almost occurred, but was prevented.

Perhaps part of the reason that Susan did not hear of the Therac-25 incidents until much later was that the hospital where she worked got rid of the Therac-4, moved their facilities, and bought a new set of linear accelerators. Susan estimated the average life of the linear accelerator to be between 5 and 10 years. After that, she said, the accelerator tends to act somewhat like an old car in which the engine light is coming on all the time. The accelerator's computer generates many faults that can become a nuisance to the operators and to the patients. Responsible operators will continue to report these faults to the physicist, when one is available, and eventually, the machine is replaced.

Susan feels that one of the biggest problems in her industry today is the lack of rigorous industry-wide standard certification and education for operators. Susan reported that there are about 102 radiation schools in the country, and that there are also different types of schools. Students are able to receive a certificate from a certificate program, usually about 12 months in length. Students are also able to receive a four-year bachelor's degree from certain schools. The American Registry of Radiologic Technologists (ARRT) provides a test that graduates of these programs may then take in order to be considered licensed entry level technicians. The ARRT also requires that therapists maintain their training through continuing education. Therapists must have 24 credits in two years before they may re-register their licenses.

In spite of the fact that the ARRT provides these guidelines for licensure, many states in the U.S. do not require hospitals or clinics to hire licensed radiation therapists. Some states require very basic exams, but, according to Susan, that in essence means that in many states anyone off the street could learn how to operate a machine, take one of these basic exams, and then be qualified to operate radiation therapy machines.

Susan and many of her colleagues continue to fight for mandatory standard certification of radiation therapists. The safety of patients depends on all of the elements of their systems of treatment working together correctly. The more operators are trained to know about the process, the more they will be able to help prevent accidents. Well-trained operators can double-check radiation dose prescriptions and question doctors when something does not seem right. With the benefit of extensive training, operators have a better sense of when it is alright to over-ride a fault message from the computer.

Well trained technicians will also be better equipped to stand up to hospital administrations that attempt to put pressure on technicians to push large numbers of patients through treatment in spite of possible dangers. Though Susan does not feel this kind of pressure from her own administration, she knows that other technicians in other clinics definitely do, especially at "free-standing" clinics that operate for profit. Susan is

aware that at these clinics there is a tremendous amount of pressure put on machine operators to get patients through treatment.

Susan also described incidents in which technicians left institutions because they didn't feel that the institutions' radiation therapy practices were safe for patients. Because there is no federal law regulating how many times an operator can re-attempt therapy after the computer displays a fault or shuts down, some operators allegedly use jumper cables that continuously override their computer's emergency shut down mechanism. Susan cited a lack of regulation, lack of training, and lack of adequate funding as reasons for these procedures.

Another issue in the radiation therapy industry that worries Susan is the fact that linear accelerator manufacturers charge large fees for operator training sessions, software upgrades, and machine maintenance contracts. When a radiation therapy machine is purchased, it comes with many binders full of information provided by the company. The clinic is given the option to buy service contracts and send physicists and operators to the company headquarters for training. Susan reported that in many clinics where money is tight, administrators are forced to choose between machine servicing contracts, software upgrades, and training.

According to Susan, mistakes are still made in the radiation therapy treatment of patients. Much of the information and calibration is still done by human beings and subject to human error. As an instructor, Susan teaches her students to anticipate every angle of the treatment, and then to check, and re-check their work. Susan also mentioned that while she teaches her students not to trust wholly in the machinery and its software, operators are largely dependent on manufacturers and hospital physicist teams to keep the machines running correctly.

Susan has a positive outlook regarding the radiation therapy industry. She knows that thousands of patients benefit greatly from radiation therapy technology. While Susan continues to push for operator certification legislation, she focuses on training her own staff well. Susan and her administration also focus heavily on quality patient care.

When asked if she thought it would be important for the designers of the software that runs the machines to know what it is like to do her job, Susan's reply was an emphatic yes, though she doubted many of the software designers of her machinery had spent much time observing a radiation treatment facility.

How a Medical Linear Accelerator Works

Generating an Electron Beam

Early radiation therapy machines used a radioactive source like cobalt to produce the ionizing radiation needed to treat cancerous tissue. Some machines still use an active radiation source. But most radiation therapy today is done with a linear accelerator. In principle, a linear accelerator works just like the computer monitor you are probably using to read this web page. The electrons are accelerated by the gun in the back of the monitor and directed at the inside of the screen, where phosphors absorb the electrons and produce light. A medical linear accelerator produces a beam of electrons about 1,000 times more powerful than the standard computer monitor. The longer a linear accelerator is, the higher the energy of the beam it can produce. The innovation of Therac 25 was that the designers found a way to fold the beam back and forth so a very long accelerator could be fit into a smaller space. Thus powerful beams could be produced, but within a reasonable amount of space

Getting the Beam into the Body

Patients can be treated directly with the resulting electron beam, as long as the beam is spread out by scanning magnets to produce a safe level of radiation. The medical linear accelerator spreads and directs the beam at the appropriate place for treatment. The picture below shows a typical medical linear accelerator in operation.

But a difficulty with the electron beam is that it diffuses rapidly in tissue and cannot reach deeper tissue for treatment. The picture in the resources section is a simulation (produced by the Stanford Linear Accelerator Center) of an electron beam traveling through air and entering human tissue. You can see the beam quickly diffuses and therefore does not penetrate deeply.

To solve this problem, Therac-25 and many other machines can switch to a mode in which X-ray photons are used for treatment. These penetrate much more deeply without harming intervening tissue. To do this, the electron beam is greatly increased in intensity and a metal foil followed by a beam "flattener" is placed in the path of the electron beam. This transforms the electron beam into an X-ray (called photons in some literature). This process is inefficient and requires a high intensity electron beam to produce enough X-ray intensity for treatment. Therac-25 used a 25 MeV electron beam to produce an X-ray for treatment. 25 MeV is 25 million electron volts (eV -- an eV is the energy needed to move one electron through a potential of one volt).

Therac-25 was what was called a dual-mode machine. It could produce the low energy electron beams for surface treatment and it could also produce a very high intensity electron beam that would be transformed into an X-ray by placing the metal foil in the path of the beam. The serious danger in a dual mode machine is that the high-energy beam might directly strike the patient if the foil and flattener were not placed in its way.

Radiation Absorbed Dose

Although MeVs are used to measure the strength of the electron beam, the measure used for therapeutic uses is the radiation absorbed dose (rad). This is a measure of the radiation that is absorbed by tissue in a treatment. Standard single radiation treatments are in the range of 200 rads. 500 rads is the accepted level of radation that, if the entire body is exposed to it, will result in the death of 50% of the cases. The unprotected electron beam in the Therac-25 is capable of producing between 15,000 and 20,000 rads in a single treatment. The unprotected beam is never aimed directly at a patient. It is either spread to a safe concentration by scanning magnets or turned into X-rays and reduced by a beam flattener.

Therac Glossary

Actuator: Device for moving the turntable
Class 1 recalls: the most serious recalls in terms of health risk
AECL: Atomic Energy Canada Limited. A Canadian company that designed and manufactured the Therac-25.
Collimator: a device for obtaining a particle beam of limited cross section
Dosimeter: radiation dose measuring device
electron beams: Accelerated electrons are absorbed by phosphors, which in turn produce light
eV: electron-volt, the energy needed to move one electron through a potential of one volt.
Gantry: the turntable assembly
GAO and Comptroller General: The General Accounting Office is the investigative arm of the Congress and is charged with examining all matters relating to the receipt anddisbursement of public funds. The General Accounting Office (GAO) was established by the Budget and Accounting Act of 1921 (31 U.S.C. 702), to independently audit Government agencies. Over the years, the Congress has expanded GAO's audit authority, added new responsibilities and duties, and strengthened GAO's ability to perform independently. The Office is under the control and direction of the Comptroller General of the United States, who is appointed by the President with the advice and consent of the Senate for a term of 15 years.
Kludge or Kluge: a computer system made up of poorly matched components
medical linear accelerator: a device that accelerates electrons to create an electron beam.
Operator: the individual responsible for the facitily room and preparing the Therac-25 machine for a particular patient.
Potentiometer: a device that independently monitors turntable position
rad (radiation absorbed dose): the amount of radiation that is absorbed by tissue in a treatment. Acceptable level for single treatment is around 200 rad.
radiation therapy: exposure to ionizing radiation using electron, X-rays or gamma rays. It is administered in a series of sessions occurring over several weeks,
X-ray (photons): High intensity electron beam (25 MeV) that is transformed. Often used to in treating deeper tissue areas.
Therac-25: a medical linear accelerator that folds the electron beam back and forth. This allows for a higher energy beam to be produced in a smaller space.
25 MeV: 25 million electron volts. This is the electron beam used in Therac 25

Analysis Documents

Socio-technical System.

The safety of the Therac-25 is not really a property of the machine alone. Accidents that go unreported contribute to (or at least fail to stop) later accidents. When the TV camera in the room is unplugged, the operator cannot see that the patient is in trouble. So safety is really a property of the entire technical and social system (socio-technical system). In a similar manner, an ethical analysis of the issues in this case requires an awareness of the entire socio-technical system.

The Therac-25 Medical Linear Accelerator is a large machine that sits in a room designed just for it. We think of the machine itself or the machine-in-the-room as the system. But the larger system, or the Socio-Technical system, that we need to think about includes:

- *Hardware*: The mechanics of the machine itself, including its associated computer
- *Software*: the operating system of the computer and the operating system of the machine
- *Physical surroundings*: the room with its shielding, cameras, locking doors, etc.
- *People*: operators, medical physicists, doctors, engineers, salespeople, managers at AECL, government regulators
- *Institutions*: AECL, FDA, each medical facility, associations of operators, etc.
- *Procedures*
- *Management models*: AECL's model of how risk is managed
- *Reporting relationships*: who was required to report accidents to whom
- *Documentation requirements*: for the software, for the facilities, for the FDA
- *Data flow*: how different parts of AECL shared information, how information was shared among agencies and organizations, how data was used by the Therac software.
- *Rules & norms*: what patients are "normally" told, what operator & physicist responsibilities are, expectations set for the programmer
- *Laws and regulations*: Reporting requirements, FDA enforcement mechanisms, medical liability law
- *Data*: data was collected in FDA approval process, use of data in Therac software.

The following list presents some of these items. We provide more information about the socio-technical system later.

- The Machine
    - Supporting Systems (video, audio, etc.)
    - Hardware
    - Software Systems
- Hospitals and Clinics
    - Doctors, Medical Physicists
    - Management, User Groups
    - Operators, Reporting Procedures
- Atomic Energy Canada, Ltd.
    - Management, Reporting Procedures,
    - Design Teams, Sales Staff, Support and Field Engineers

- Government Medical Device Regulation
    - Food and Drug Administration
    - Canadian Radiation Protection Bureau
    - Reporting Procedures

A thorough investigation of the Therac-25 case requires some grasp of most of these items. You will come across most of these items as you read this case. Setting your sights on the entire system will help you avoid the trap of finding a single point of blame. It is easy, for instance, to decide that the programmer made serious mistakes and to end one's analysis there. This is a short-sighted approach. It would miss the problems with maintenance in the cancer therapy facilities; it would miss the incomplete reporting requirements for the FDA; it would miss the inadequate and misleading testing of the Therac-25 system.

The machine and software.

There were two previous versions of Therac machines, each produced by AECL in collaboration with a French company, CGR. Therac 6 and Therac 20 (each named for the power of the beam they could produce) were based on earlier designs from CGR. By the time Therac-25 was released for sale, AECL had 13 years of experience with production of medical linear accelerators. Therac-25 was based on these previous versions. Its main innovations were (1) a "double pass" electron beam so the machine could produce more energy in less space, and (2) the addition of extensive computer control of the machine. This latter innovation allowed AECL to move much of the checking for hazardous conditions into the software.

The Therac-25's ancestors, Therac-20 and Therac-6, had used a minicomputer (a DEC PDP-11) to add some convenience to the standard hardware of a medical linear accelerator. They both could work without computer control. AECL determined to make its new model, Therac-25, a tightly-coupled combination of software and hardware. Therac-25 software was not written from scratch, but was built up from components that were borrowed from the earlier versions of Therac.

Therac-25 was a dual mode machine. This means that it could treat the patient with relatively low energy electron beams or with X-ray beams. This dual mode allowed for further cost savings in that two machines could be replaced by one. Therac-25 also had a "field light" position that allowed a standard light beam to shine in the path of treatment to help the operator in setting up the machine. Thus there were three modes in which the Therac-25 could operate: electron beam and X-ray for treatment, and field light for setup.

Even though they are relatively low energy, the electron beams are too powerful in their raw form to treat the patient. They need to be spread thinly enough to be the right level of energy. To do this, Therac-25 placed what are called scanning magnets in the way of the beam. The spread of the beam (and thus its power) could be controlled by the magnetic fields generated by these magnets. Thus for electron beam therapy, the scanning magnets needed to be placed in the path of the beam. It was a race condition produced by a software error in setting the magnets and resulting in a turntable mismatch that produced at least two of the accidents.

X-ray treatment requires a very high intensity electron beam (25 MeV) to strike a metal foil. The foil then emits X-rays (photons). This X-ray beam is then "flattened" by a

device below the foil, and the X-ray beam of an appropriate intensity is then directed to the patient. Thus, X-ray therapy requires the foil and the flattener to be placed in the path of the electron beam.

The final mode of operation for Therac-25 is not a treatment mode at all. It is merely a light that illuminates the field on the surface of the patient's body that will be treated with one of the treatment beams. This "field light" required placing a mirror in place to guide the light in a path approximating the treatment beam's path. This allowed accurate setup of the machine before treatment. Thus, for field light setup, the mirror needed to be placed in the path where one of the treatment beams would eventually go.

In order to get each of these three assemblies (scanning magnets or X-ray target or field light mirror) in the right place at the right time, the Therac-25 designer placed them on a turntable. As the name suggests, this is a rotating assembly that has the items for each mode placed on it. The turntable is rotated to the correct position before the beam is started up. This is a crucial piece of the Therac-25 machine, since incorrect matching of the turntable and the mode of operation (e.g. scanning magnets in place but electron beam turned on high for X-ray) could produce potentially fatal levels of radiation. The original Leveson and Turner (ref) article includes diagrams of the machine and the turntable, and does the website.

Setup and Actuation. The Therac-25 operator sets up the patient on the table using the field light to target the beam. In doing this, treatment parameters must be entered into the machine directly in the treatment room. He or she then leaves the room and uses the computer console to confirm the treatment parameters (electron or X-ray mode, intensity, duration, etc.). The parameters initially entered in the treatment room appear on the console and the operator simply presses return to confirm each one.

The computer then makes the appropriate adjustments in the machine (moving the turntable, setting the scanning magnets, setting beam intensity etc.). This takes several seconds to do. If the operator notices an error in the input parameters, he or she can, during the setup, edit the parameters at the console without having to start all over again from inside the treatment room. It was a race condition produced by editing of parameters and resulting in a mismatch of energy with turntable position that produced at least two of the accidents.

When the computer indicates that the setup has been done correctly, the operator presses the actuation switch. The computer turns the beam on and the treatment begins. The main tasks for which the software is responsible include:

- Monitoring input and editing changes from an operator
- Updating the operator's screen to show current status of machine
- Printing in response to an operator commands
- monitoring the machine status
- rotating the turntable to correct placement
- strength and shape of beam
- operation of bending and scanning magnets
- setting the machine up for the specified treatment
- turning the beam on
- turning the beam off (after treatment, on operator command, or if a malfunction is detected)

The Therac-25 software is designed as a real-time system and implemented in machine language (a low level and difficult to read language). The software segregated the tasks above into critical tasks (e.g. setup and operation of the beam) and non-critical tasks (e.g. monitoring the keyboard). A scheduler handled the allocation of computer time to all the processes except those handled on an interrupt basis (e.g. the computer clock and handling of computer-hardware-generated errors).

The difficulty with this kind of software is the handling of things that might be occurring simultaneously. For example, the computer might be setting the magnets for a particular treatment already entered (which can take 8 seconds) while the operator has changed some of the parameters on the console screen. If this change is not detected appropriately, it may only affect the portion of the software that handles beam intensity, while the portion of the software that checks turntable position is left thinking that the old treatment parameters are still in effect. These sorts of scheduling problems when more than one process is running concurrently are called race conditions and are the primary problem that produced the accidents.

In 1983, just after AECL made the Therac-25 commercially available, AECL performed a safety analysis of the machine using Fault Tree Analysis. This involves calculating the probabilities of the occurrence of varying hazards (e.g. an overdose) by specifying which causes of the hazard must jointly occur in order to produce the hazard.

In order for this analysis to work as a Safety Analysis, one must first specify the hazards (not always easy), and then be able to specify the all possible causal sequences in the system that could produce them. It is certainly a useful exercise, since it allows easy identification of single-point-of-failure items and the identification of items whose failure can produce the hazard in multiple ways. Concentrating on items like these is a good way to begin reducing the probabilities of a hazard occurring.

In addition, if one knows the specific probabilities of all the contributing events, one can produce a reasonable estimate of the probability of the hazard occurring. This quantitative use of Fault Tree Analysis is fraught with difficulties and temptations, as AECL's approach shows.

In order to be useful, a Fault Tree Analysis needs to specify all the likely events that could contribute to producing a hazard. Unfortunately, AECL's analysis left out consideration of the software in the system almost entirely. Since much of the software had been taken from the Therac-6 and Therac-20 systems, and since these software systems had been running many years without detectable errors, the analysts assumed there were no design problems in the software. The analysts considered software failures like "computer selects wrong mode" but assigned them probabilities like 4×10^{-9}.

These sorts of probabilities are likely assigned based on the remote possibility of random errors produced by things like electromagnetic noise, or perhaps the mean-time-between-failures data generally available then for PDP-11 machines. They do not at all take into account the possibility of design flaws in the software. This shows a major difficulty with Fault Tree Analysis as it was practiced by AECL. If the only items considered are "failure" items (e.g. wear, fatigue, etc.) a Fault Tree Analysis really only gives one a reliability for the system.

Hospitals.

The complexity of cancer treatment organizations is one of the things students must deal with as they struggle to understand why the accidents happened and to construct a plan to respond to the accidents.

Cancer treatment facilities are often housed in large hospitals, but some are stand-alone cancer treatment centers (like the Tyler, Texas center). Those associated with hospitals are more likely to be non-profit, while those that stand alone are more likely to be for-profit organizations. Financial pressures are likely to be strong at both for-profit and not-for-profit organizations, but they will have slightly different regulatory structures.

During the time of Therac-25 (the mid 80s) a well equipped treatment facility might have 3 different machines. The machines would be capable of producing different kinds of radiation, different strengths of beam, and capable of different kinds of exposure to the patient. Each of these machines would cost, for the machine alone, between 1 and 2 million dollars. In addition, special housing for each machine is needed, with shielding in the walls, adequate power supply, video and intercom links, etc.

Operators would be needed to run each machine. For larger facilities, a supervisor of the operators, with more training and experience might be needed. In addition, at least one MD specialist in cancer radiation therapy (a Radiation Oncologist) would be required. Finally, a medical physicist would be needed to maintain and check the machines regularly. Some facilities contract out the services of a medical physicist. Finally, all the support personnel for these specialists (nurses, secretaries, administrative staff, people to handle billing and paperwork, janitorial staff, etc.) are required.

Medical Linear Accelerators do age over time, and older machines often produce more errors. Five to ten years is a reasonable life span for a machine. Thus, simply to maintain a set of three medical linear accelerators, an institution can expect to spend 1 to 2 million dollars every third year.

Sometimes errors can be resolved and a machine kept longer using software upgrades or upgrades or retrofits of machine parts. The companies that sell linear accelerators charge maintenance contracts that can include different levels of support. Because of monetary constraints, sometimes facilities are forced to choose between software updates, manuals, and training for operators and physicists. All this is in addition to the price of the machine itself.

Production pressures are always present when an expensive medical technology is being used. These very expensive machines need to treat enough patients to pay for themselves over their lifetime. And in for-profit medical facilities the additional pressure of generating a profit is added to this production pressure. Another kind of production pressure is generated because of concern for the patient. Patients' schedules require treatments on certain days and it disrupts the patients' lives and slows down their treatment to have to reschedule them for another day while the machine is being checked out.

These production pressures generate the desire to "push patients through." If a machine gives only a portion of the prescribed dose, an operator will often repeat the treatment with enough radiation to add up to the total prescribed dose. Of course, because of liability issues and concerns for patient welfare, this can only be done when it is thought safe.

One of the advantages of the significant computerization of the Therac 25 machine was that setup for treatment could be done much more quickly. This allowed the operator more time to speak with the patient and interact with them about their health concerns. In addition, this increased efficiency allowed more patients to be scheduled during a day. Thus, more patients could be treated, but the atmosphere was not reduced to that of a factory.

Facilities that run medical linear accelerator are surely concerned about liability for injury to patients that might occur. Insurance, for medical providers, is quite expensive and errors in treatment can result in lawsuits, which in turn produce increases in insurance premiums. Standard practice in litigation is to "sue everyone with deep pockets." This means that even if an error is the result of poor design of a linear accelerator, the facility itself will be sued simply because they were involved: they have insurance and thus "deep pockets."

But it is in the interest of facilities to reduce errors without the threat of lawsuits. When a treatment must be restarted several times because of errors, it may reduce patient confidence in the facility. This can mean patients moving to another facility with which they are more comfortable. Finally, medical professionals are in their business because they want to help people and have the knowledge and skill to do so. So a primary motivation of medical professionals is patient welfare.

The Food and Drug Administration.

In addition to dealing with the technical issues and organizational issues associated with the hospitals, students need to consider the role the FDA plays in regulating medical devices. Understanding the constraints the FDA imposes on possible solutions (and the opportunities they provide) is a crucial part of designing a responsible solution to the accidents.

The Food and Drug Administration (FDA) was created when Congress passed the Food and Drugs Act in 1906. This act was the first of a series of laws and amendments that gave the FDA jurisdiction over the regulation of foods and patent medicines. In 1938, Congress strengthened and expanded the FDA, to include the regulation of therapeutic and medical devices within its jurisdiction.

The FDA's Bureau of Medical Devices and Diagnostic Products was created in 1974, and soon operated in conjunction with the Medical Devices Amendments of 1976. The amendments helped to clarify the logistics of the regulation of medical devices, and required the FDA to "ensure their safety and effectiveness."

Radiation had been recognized as a health hazard since before World War I, and the FDA monitored the health risks that radiation emitting products posed to America's workers and consumers. As FDA's responsibilities for monitoring radiological devices grew, a bureau within the FDA called the Center for Devices and Radiological Health (CDRH) was established.

In 1980 the FDA's budget had swelled to over $320 million, with a staff of over 7,000. Many bureaus controlled areas such as biological drugs, consumer products, public health standards, and veterinary medicines.

FDA approved medical devices before they "went to market." This was called Pre-Market Approval and was a somewhat complex process. In the FDA Pre-market

Approval scheme, devices were organized into three classes, as established by the 1976 Medical Device Amendments.

- Class I devices, "general controls provide reasonable assurance of safety and effectiveness," for example bedpans and tongue depressors
- Class II devices, such as syringes and hearing aids, "require performance standards in addition to general controls"
- Class III devices like heart valves and pacemakers are required to undergo pre-market approval as well as complying with general controls
- In addition to classifying devices as Class I, II, or III, FDA approved devices for market in one of two ways:
    - Proof of *Pre-market Equivalence* to another device on the market, termed 501(k)
    - OR *Pre-market Approval* (Rigorous Testing)

If a company could show Pre-market Equivalence (proof that a new product was equivalent to one already on the market), the new product could be approved by FDA without extensive, costly, rigorous testing. In 1984 about 94% of medical devices came to market through Pre-market Equivalence.

If a product was not equivalent to one that was already on the market, FDA required that the product go through testing to gain Pre-market Approval. In 1984 only about 6% of medical devices were required to go through this testing.

Thus, it was clearly in the interest of medical device producers to show that their product had pre-market equivalence. The Therac-25, brought to market in 1983, was classified as a Class III medical device. Since AECL designed the Therac-25 software based on software used in the earlier Therac-20 and Therac-6 models, Therac-25 was approved by FDA under Pre-market Equivalency. This declaration of pre-market equivalence seems optimistic in that (1) most of the safety mechanisms were moved into the software, a major change from previous version of the machine, and (2) the confidence in the safety of much of the software was based on its performance in the older machines, which had hardware safety devices installed to block potential accidents.

A 1983 General Accounting Office (GAO) report criticized the FDA's "adverse experience warning system" as inadequate. FDA had published reports about potential hazards, including reports in their own newsletter, *The FDA Consumer*. The FDA implemented the mandatory medical-device reporting rule after Congress passed the Medical Device Reporting Legislation in 1984. This rule required manufacturers to report injuries and problems that could cause injuries or death.

Before 1986, users of medical devices (hospitals, doctors, independent facilities) were not required to report problems with medical devices. Instead, under the medical device reporting rule, manufacturers of these devices were required to report problems. The idea was that manufacturers would be the first to hear about any problems with the devices they made and that therefore reports would be timely. In addition, manufacturers would be most likely to have the correct information needed about a device to help resolve difficulties.

In the mid-1980s, the FDA's main enforcement tools for medical devices already on the market were publicity. The FDA could not force a recall, it could only recommend one. The CDRH (Center for Devices and Radiological Health monitors radiological devices) issues its public warnings and advisories in the *Radiological Health Bulletin*.

Before issuing a public warning or advisory, the FDA could negotiate with manufacturers in private (and in the case of Therac 25, with regulatory agencies in Canada). In response to reports of problems with a medical device, the FDA could, in increasing order of severity:

1. Ask for information from a manufacturer.
2. Require a report from the manufacturer.
3. Declare a product defective and require a corrective action plan (CAP).
4. Publicly recommend that routine use of the system on patients be discontinued.
5. Publicly recommend a recall.

Thus, even when the FDA became aware of the problem, they did not have the power to recall Therac-25, only to recommend a recall. After the Therac-25 deaths occurred, the FDA issued an article in the *Radiological Health Bulletin* (Dec. 1986) explaining the mechanical failures of Therac-25 and explaining that "FDA had now declared the Therac-25 defective, and must approve the company's corrective action program."

After another Therac-25 overdose occurred in Washington state, the FDA took stronger action by "recommending that routine use of the system on patients be discontinued until a corrective plan had been approved and implemented"[2] (. AECL was expected to notify Therac-25 users of the problem, and of FDA's recommendations.

After the Therac-25 deaths, the FDA made a number of adjustments to its policies in an attempt to address the breakdowns in communication and product approval. In 1990, health- care facilities were required by law to report incidents to both the manufacturer and FDA.

AECL and the state of the technical art.

A crucial part of this case is understanding problems with synchronization of concurrent processes (explained in detail in section X). But we also need to understand what the programmers of the Therac-6, Therac-20, and Therac-25 were likely to have known about these issues. Almost nothing is known about the specific qualifications of the Therac-25 programmers (or even their identities), but we can get some idea of the current state-of the art at the time.

Although software solutions to synchronization problems were known in the mid-1970s when AECL and CGR developed the Therac-6 software, it seems unlikely that those programmers would have used them in their implementation. An elaborate and complicated solution by Dekker[3] was available, but was "a tremendous mystification,"[4] difficult to comprehend or to program correctly. Strategies that depend on adding special features to the operating system appeared beginning with Dijkstra[5], but such operations did not appear as a standard part of common operating systems until years later, and their implementation in the Therac-6 system seems unlikely unless they had a specialist on their team. Operating systems courses at the time focused on theoretical discussions

[2] Radiological Health Bulletin, March 1987
[3] Cited on p. 58 of Dijkstra, E.W., "Co-operating Sequential Processes," in Genuys, F., ed., *Programming Languages,* Academic Press, 1965
[4] Ibid., p.66
[5] Ibid.

rather than practical implementation issues. John Lions' famous commentary on UNIX Version 6 source code (Lions, 1996), developed in 1975-76 for his students in Australia, is widely considered as the first operating systems course text to address practical concerns seriously (Tanenbaum, 1987).

Thus, assembly programmers in the mid-1970s would undoubtedly have employed hardware solutions for synchronization, if they were even aware of subtle issues such as race conditions. A "test and set lock" (TSL) instruction provided one approach, in which a particular machine instruction had the capability to copy a shared variable's value to another location held privately by a task, and to assign a value to that shared variable, all in a single indivisible operation. Other machine instructions (e.g., SWAP) could serve the same purpose. However, an examination of the PDP-11 instruction set[6] shows that no such instruction exists on the machine used for the Therac systems. It is conceivable that the "subtract one and branch" (SOB) instruction, designed for loop control, might have been turned to this purpose by a creative and careful programmer who had awareness of these synchronization issue.

These facts hardly exonerate the Therac programmers. The operating-system level routines for Therac-25 were written specifically for that system according to Leveson and Turner; those programmers had a responsibility to know about issues such as race conditions in multitasking real-time systems. They would most likely have heard about postponements of the release of the IBM 360 time-sharing system and of Multics. Both of these projects were late in part because of the difficulty of getting synchronization done properly. However, unless those responsible for the operating-system executive routines had prior experience writing concurrent software, it seems quite conceivable that they had never seen the subtleties of race conditions.

[6] Available at http://www.village.org/pdp11/faq.pages/PDPinst.html

Ethical Reflections.

In Therac-25, safety is the primary concern. But we need to think about issues of safety at several social levels. You will recognize this as the analysis approach used in chapter 2. Each of these levels makes evident different issues and different ethical concerns in the case. In addition, the grid reminds us that safety issues are only one of many issues that might concern us.

Quality of Life

The decision to computerize a medical linear accelerator was, in the beginning, a quality of life consideration. AECL did not set out to make a device that would expose individuals to harm. It made improvements in a device that would, in theory, allow greater access to a medical technology and would increase the quality of care those patients received. Operators were pleased with the new interface and felt it gave them more time to interact with the patient as a real person rather than merely being a technician to the machine. Hospitals were pleased that the machine, being dual-mode, could offer a wide variety of treatments for a substantially reduced capital investment. Because the machine was less expensive, patients could have access to a treatment that they might not have been able to afford. Thus, the machine's design made significant improvements in quality of life for many people.

But in the process, a lax culture of safety in the organization led to a system design that was unsafe and not well tested. In this case, the values of safety and increased quality of life for consumers *need not* have been in conflict. But in practice, they became so.

Power

Perrow (ref) suggests that most risk analysis procedures are really a way for *some* people to think clearly about the risks to which they will subject *other* people. People who are doing a risk or safety analysis are usually those hired by the company to protect itself from risk. There are mixed motives here: by protecting themselves from risk, they also protect the safety of those using their products.

Ford Motor Company made itself infamous by explicitly comparing risk to the company (in dollars lost from lawsuits) to risk that consumers faced (from inadequate design of gas tanks in the Pinto). They decided that it would cost less to pay the lawsuits than to fix the car. Here the calculations were all financial. But it is at least up for debate whether all companies make decisions in this manner. In many, the motives are mixed: protection of the company and safety of the consumer.

But AECL's priority seems odd even in the light of self-protection. Its risk analysis seemingly was not done to protect the company, but to certify their already strongly held belief that the machine was safe. This sort of unfounded optimism regarding technology at least provides them with the defense of ignorance. But this defense is less persuasive when offered by those with power over other's well-being. Often, when individuals or corporations are given more power, we are also more likely to hold them more responsible for their actions.

At any rate, this case is clearly an issue of who has power to enforce the acceptance of risks on others. This power may be economic (as in the case of AECL and the hospitals) or political (as in the FDA).

But at the individual level, the power may simply be role-based—acquired because you happen to be the software engineer assigned to a particular project. This is what Huff (ref) has called unintentional power—the power that a designer has over the users of a product. Someone with unintentional power uses it without intending benefit or harm to the ultimate user of the product. This is another case of the defense of ignorance. In this case, the defense is harder to believe, since the software controlled a potentially lethal radiation beam. But the intention of the programmer, or of the operator, was not to harm patients, or even to place themselves in the position where they could harm them. And the discussion on race conditions in the resources section makes the case that this may have been a relatively unknown problem with real-time systems during Therac-25's design. Still, taking a job as a software engineer entails this unintentional, positional power. It is better to know this than to ignore it.

These sorts of power differentials exist at all levels of the social analytical framework. And a careful ethical analysis of power will ask what duties go along with that power, and what rights are held by those with less power.

System Safety

We cannot present here a full analysis of system safety and instead refer the interested reader to items in our bibliography and to various web sites that address these issues. You can see what noted system safety expert Nancy Leveson has to say about this case in the excerpt from her article that we provide in the supporting documentation.

However, we have at least learned that in order to understand the safety issues properly, we must look at them at several levels of social complexity, just as the ImpactCS framework suggests.

Safety at the Individual Level

The programmer. Certainly the single individual who did the programming for Therac-25 had responsibilities as a computing professional. To whom were these responsibilities owed? An obvious first responsibility is to the organization that employed him. Another party to whom responsibility is owed is to the eventual users of the linear accelerator: the patients. We can certainly add to these two (e.g. to the profession, to the machine operators), but let's take each of these for the purposes of this analysis.

The programmer's responsibilities to his employer were more than simply to do as directed to by the other designers of the system (or by whoever his immediate superiors were). He had a responsibility to make his superiors aware of the dangers inherent in doing safety interlocks only in the software. Whether this danger was obvious to him or not is an interesting question. Even today many computing professionals place more confidence in the safety of software than is likely. Software safety was little understood

at the time the Therac-25 system was designed (see the section on race conditions in resources).

The point here is that a computer professional is responsible to the employer for using the best available methods to solve the software problems with which he or she is confronted. There are a variety of professional decisions the programmer made in the Therac-25 design that suggest he was lax in this responsibility (e.g., using unprotected memory, improper initialization, lack of appropriate testing, etc.). Thus, as an employee, he fell short of the mark in providing his employer with professional work. We cannot know whether this shortcoming was one of lack of knowledge or resources, poor execution, time pressure, or some other extenuating circumstance.

In addition to responsibilities to his employer, the programmer clearly had a responsibility to the users of the technology he was designing. In the context of safety, his responsibility was to design software that minimized the likelihood of harm from a dangerous medical device. This obligation to "do no harm" need not mean that software should never be paired with medical linear accelerators. From the perspective of the operator we interviewed this pairing was a positive benefit in making setup and treatment easier. But it does mean that, to the extent it was within the professional control of the programmer, he should have designed the system to do no harm while providing this positive good. Again, whether the failure to do this was a result of a lack of knowledge or of poor execution, we cannot know.

To sum up, the programmer had clear responsibilities to both his employer and to the users of the device. He clearly failed in these responsibilities. If we were interested in blame, we could not tell the amount of blame to assign here. We know nothing of the programmer's background or training. Thus we cannot know if the programmer knew how poor the software design and testing was. Nor do we know much about the software routines that the programmer reused to put together the system for Therac-25. Nor do we have any idea of the internal company dynamics that may have resulted in the lack of testing.

The operators. We noted in the case narrative that the operators of the linear accelerators had a complex combination of responsibilities. Chief among these are responsibilities to their employer and to the patients.

Just like the programmer of the system, we know little about the background and training of the operators in this case. But we can at least specify their responsibilities. They were responsible to their employers to operate the machine efficiently, getting all the scheduled treatment done in any particular day. They also had a responsibility to their employer to look after the machine so it could be maintained properly. Finally they had a responsibility to their employer to operate the machine carefully and not to place patients in danger.

From published accounts of operator's comments, from our interview of a Therac-4 operator, and from comment's made in court documents, it seems clear that none of the

operators felt they were placing their patients in any danger when they pressed the button. Thus, we can rule out intentional negligence. But what happened? Leveson (ref) suggests that the interface on the console made operators tolerant of error messages, and readily rewarded them for pressing the "proceed" button whenever a minor error appeared. The interface made no distinction between life threatening errors and minor errors, except that major errors would not allow a "proceed." Given this, it is hard to see how the operators might be responsible for the errors, even though they were the ones to press the key.

An interesting issue arises because of the current move among operators to become more professionalized. As operators are better trained, are certified, and are more aware of the workings of the machine, they gain prestige—but they also gain responsibility. As they become well trained enough to foresee such errors, their responsibility for them will increase.

Safety at the Group Level
There are two organizations at this level whose actions need to be thought about: the treatment facilities and AECL.

Atomic Energy Canada Limited. With regard to safety in this case, AECL's responsibilities in making a medical linear accelerator are to a range of individuals: their shareholders, their employees, the governments of Canada and the United States, to the facilities that bought the machine, and finally to the patients who were treated by them. Responsibility to shareholders and employees are similar, and for this analysis will be considered the same.

Before we look at these specific responsibilities, we will need to understand some of the technical issues involved in the analysis of a system for safety. In this instance, technical knowledge is required to make ethical judgments.

AECL claimed to do a safety analysis of its machine, but in fact the analysis only shows the likelihood of the system failing because a part wears out. There was apparently no systematic search for design flaws in the software until after the FDA required an analysis. Unfortunately, a system can be highly reliable but thereby reliably kill people because of a design flaw. This confusion of reliability analysis and safety analysis is a critical failing on the part of AECL.

Some indication of the motivations behind AECL's inadequate safety analyses can be gleaned from the way AECL appeared to use probabilities in its analysis. These probabilities seemed to be assigned to quantify and to prove the safety of the system, rather than to identify design flaws. For example, after redesigning the logic to track the microswitches that indicated the position of the turntable, AECL apparently used a sort of Fault Tree Analysis to assert that the safety of the system had been improved by at least 5 orders of magnitude. This astonishing claim of improvement is applied to the safety of the entire machine. This use of probabilities from a Fault Tree Analysis can effectively hide critical design flaws by inflating the perception of reliability and discouraging

additional search for design flaws. This hiding of design flaws was a tragic, if unintentional side effect of the improper use of this analysis.

Thus, in failing to look systematically for design flaws in its software, AECL left itself (and its employees and shareholders) open to liability claims from injured consumers. This is clearly also a failure of its responsibility to patients and to the facilities who bought the Therac-25 machine and who were assured there was no way it could hurt patients. This failure must be perceived in the light of prevailing standards (or lack thereof) in system safety at the time of the design and release of Therac-25.

The Cancer Treatment Facilities. The cancer treatment centers were primarily consumers of a product that is tested, maintained, and certified by others. This product was sold to them with assurances that it could not hurt patients (see Honesty & Deception). And the facilities do not have the responsibility or the capability to independently check these systems for safety.

But they do have responsibility for the safe operation and low level maintenance of the machines once they are in operation. It was clear that at least one facility fell down in this respect. In the first Tyler accident, the video monitor to the room was unplugged and the intercom was out of order. This would not have been a problem if there were no accidents—but there were. One difficulty with the safe operations of systems is that standard maintenance can become tedious and not seem a necessary component in the safe operation of a regularly used system. In this case, an individual might have been spared a second overdose if the basic communication systems had been working.

We should note, however, the extraordinary efforts of the medical physicist at Tyler in determining the cause of the overdose. This individual effort was supported by the Tyler facility and made possible by the facility's decision to have a full time physicist on staff. Some evidence of this support comes from the facility's decision to report the accident to the FDA even though there was no requirement that they do so. Note that the facility decided that their responsibility extended beyond the requirements of the law.

Thus, most facilities had relatively minimal responsibilities in this case and most seemed to fulfill them. The facilities had little power to resolve the problem and depended on AECL and on the FDA's approval process to protect them and their patients. Perhaps in this dependence they were too optimistic, but it is difficult to see what other choices they might have had.

Safety at the National Level

At the time of the Therac-25 accidents, the *Center for Devices and Radiological Health* (CDRH) of the FDA was responsible for the oversight of the immense market in radiation based therapy and diagnostics. As we have seen, most (94% in 1984) devices for the market were approved by "pre-market equivalence" and thus not subjected to stringent testing. The CDRH could not have handled the load of testing all these devices.

Since the rules for FDA are set by congress, FDA's rules need to be analyzed from the perspective of the responsibilities of congress. But FDA implementation of those rules is under its control. Thus we can ask if the CDRH (as a center in FDA) should have allowed Therac-25 to be approved under pre-market equivalence. Without more information this is difficult to determine. The CDRH did seem vigorously to follow the case once it became aware of the Tyler accidents, though there is some evidence that they were reluctant to halt quickly the use of the Therac-25 when the problems became evident. This reluctance may be because of their responsibility to not place an undue burden on manufacturers in their caution regarding a product. Or they might have had in mind the many people who were being helped by the treatment. This tension between responsibilities to manufacturers/industry and responsibilities to patients is always present in decisions by the FDA. Hindsight makes this one seems easy to decide.

One of the problems in this case is that the FDA depended on AECL to notify it of accidents that had occurred. They did not hear directly from the hospitals when the accidents happened. AECL had, at best, a mixed record of notifying the FDA of problems. So perhaps facilities should have been required to directly report accidents (they are today). But FDA could not make this requirement; it could only enforce existing law. Thus, perhaps it was a responsibility of congress to enact this law. The counter-argument is that congress should allow the market to work out these issues. But in this instance, at least, the market was too slow to save the individuals who were killed or injured.

The entanglement of different ethical issues become very clear at this level of analysis. Different constituencies will value different things (e.g. personal privacy vs. business freedom). These choices among different values are as severe at the other levels (e.g. operator's responsibility to employer and patient) but not as easily seen to the outside observer. Choice and balance among these values becomes inescapable, however, at the political level.

Safety at the Global Level

Communication between the *Canadian Radiation Protection Board* and the FDA seemed to work pretty well in this case. These two agencies had responsibilities to their respective governments, to industry in their countries, and to patients in their countries. AECL's communication with FDA did not seem to be hampered by its international flavor. However, this is a case of two relatively similar countries and cultures interacting with each other. Similarities in legal standards and in government oversight made this case easier. This might be an even less happy story if we had been dealing with widely different cultures of business or legal systems in the two countries.

Privacy

Privacy issues may well be raised by this case as one begins to recommend some sort of a national reporting mechanism for medical devices. In order to accurately report on any accident, sensitive medical data about individuals would need to be collected by treatment facilities and made available to national agencies. These national agencies

might, in turn, make this data available internationally. In our case, data about the accidents was shared by agencies of the Canadian and US governments.

It seems possible to make the data on patient records related to medical accidents anonymous — to separate the data from the identity of the patient. But one would have to think carefully about how to do this. The quickest solution, attaching the patient's medical history, is likely collecting too much data and violating the privacy of the patient. Collecting only as much data as is needed is a reasonable requirement. Deciding what is needed will be more difficult.

Property

At one point in its interactions with the user groups, AECLfound itself being asked for the source code used in the Therac 25 software. AECL claimed that it had proprietary rights to the software and would not make it public. This is another case of two values coming into conflict. A concern for safety suggests that it would be helpful to open the source code to inspection by the FDA or its agents or by the user groups. But to force this openness would violate the property rights of the owner of the software, AECL. One suspects that AECL's refusal to open the code to inspection is a defensive move based on avoiding liability rather than an attempt to protect the value of the intellectual property. But if close inspection showed the software to be poorly designed, the value of the software would surely diminish. Is this a case in which we want to uphold property rights? There may in fact be some case here for an "open software" standard to protect public safety.

The framework suggests we identify several levels of social analysis for each ethical issue. In this case we have interaction among those levels. The public (ideally, represented by the FDA) was placed at great risk by the software. What claims are there that the cancer treatment facilities and the FDA can make regarding the value of their being able to inspect the design and logic of the software? What claims can patients (or their surviving families) make regarding the validity of the claim by AECL to keeps its software a trade secret? So, in order to understand the property issue correctly, we need to look at claims made at many levels: national, organizational, and individual.

Equity and Access

Equity and access issues are also raised by this case. The whole point of the design of Therac-25 was to make a medical linear accelerator that was less expensive to produce and thus less expensive (and more available) to consumers. This is often an effect of the free market on the price of technology. If AECL could make a less expensive, but equally useful linear accelerator, it would sell more of them and they would be more easily available to the public. AECL would, doubtless, make money in the process. This is Adam Smith's invisible hand (ref) at work: decisions to make a better, less expensive product are good both for the manufacturer and for the consumer.

On the other hand, increasing regulation and oversight will impose increasing costs on providers of medical devices. These regulations may be seen as necessary, given the track

record of companies like AECL. But they still increase the development costs to the company and the cost of the product. This, in turn, reduces the availability of the devices.

Again, we have here an issue of balance between competing goods: safety of the consumer and the increased access of the consumer to life-saving medical technology.

Honesty and Deception

Honesty and Deception issues are central to this case. In several places, AECL representatives made claims of safety for the Therac-25 device that in retrospect seem at least exaggerated.

How did this occur? These claims were made by individuals (salespersons, engineers), on the behalf of an organization (AECL). Individuals making claims like these have some responsibility to check on their accuracy. But salespeople have little expertise to evaluate this information and are thus more dependent on the organization. Engineers, including software engineers, have the capability of evaluating the claims though they may be allowed little time in which to do so. Again, we find a balance between the engineer's responsibilities to the company (use time efficiently) and to the consumer (evaluate carefully claims made about the product). Because of their special expertise, it is precisely the role of a professional to balance these conflicting responsibilities, and not to neglect responsibility to the consumer.

What organizational responsibilities might there be regarding claims of safety in medical devices? AECL representatives in several instances made claims that no overdoses had occurred with the Therac-25 machine, when there was clear evidence that someone at AECL must have heard of several previous accidents. This suggests (if we are charitable) that there may have been some internal miscommunication within AECL. Some portions of the organization may have known about the lawsuit regarding radiation harm but not have had the time or seen the need to inform other parts of the organization. For instance, those in the legal division, hearing of the lawsuit, may have assumed that the engineers were aware of the issue and that there was no immediate need to contact them. This sort of miscommunication is a daily matter in organizations, even small ones (think of the miscommunication that occurs in your family).

Thus one part of the organization may have been making claims that no accidents similar to the reported ones had occurred based on the information available to them. Still, when the stakes are as high as they were in this case, organizations have a special responsibility to transmit safety critical information as quickly and as accurately as possible. This occurred sometimes within AECL (the FDA was notified quickly of the Hamilton accident) but not all the time.

•Intermediate Concepts

Professional responsibility for the quality of software

A central issue in this case is: who is responsible for the safety of software? An important distinction in American law is this instance is that between a product and a service. These are covered by different legal approaches, as we explain below.

Products are covered by *strict liability* Strict liability means that if the product causes harm, there is no need to show foreseeability to avoid the harm on the part of the seller. This sets a high standard for safety. The reasoning is utilitarian: it includes the cost of lawsuits in product prices, makes the entity with the most control over the product liable

Services are covered by *negligence*. This means that the standard is what a reasonable professional would have done. Services are rendered individually, and both the vendor and client have some control over the quality of the product.

Mass Marketed Software might best be treated as a product (and has in some legal cases). Custom software is likely best treated as a service. Certainly there are mixed cases. Still, remember that this discussion if about the *legal* issues surrounding software liability, rather than the ethical ones. It is at least reasonable to make the case that designers of custom software should take extra-ordinary care when designing software that mat be life-threatening.

What is a reasonable professional and what are prudent standards in safety-critical software? One answer to this is not to wait for the litigation in court and to set standards for licensing software engineers that involves standards for the design of safety-critical software. Fir instance, Texas provides for the licensing of Software Engineers and provides standards forprofessional practice and testing to certify that engineers meet those standards. On the other hand, the ACM has produced a code of software engineering ethics (that address issues of safety), but has rejected a proposal to endorse licensing of SE. Their concern is that standards of "reasonable professional care" are still shifting to quickly to codify them in the way it would be needed for licensing. So, professional licensing is still an issue in dispute.

One answer to the issue of software liability has been proposed in the Uniform Computer Information Transactions Act, or UCITA. Two states have adopted these legal standards (Maryland and Virginia) but they have been blocked in many other states. The ACM and IEEE have publicly argued against UCITA. The central approach of UCITA is that in enshrine licensing of software as the standard for ownership, control, and liability for software. UCITA structures the concept of software license to

- prohibit "public criticism" of software (e.g. benchmarking)
- prohibit reverse engineering
- allow "self help" to enforce licenses (e.g. remote disabling of software)
- limit resale of software
- allow modification of license by posting on a web site
- force liability lawsuits into arbitration, thus limiting damages
- limit in many other ways the remedies that users can seek for faulty software

As of this writing (2005) it looks like UCITA is stalled in adoption by any other states.

•Exercises

•Moral Problem Solving Exercises.

Procedures for doing each of the following problem solving exercise are presented in the toolbox.

Generating the Socio-technical system. The socio-technical system presented in the analysis section is certainly incomplete. When generating your STS, be sure to consider the international aspects of the situation.

Value Conflicts in the STS. The values, goals, and agendas of the FDA, the treatment facilities, AECL, and the patients are certainly different. But be sure to recognize where they re in agreement (e.g. all are interested in safety and are willing to take some risk).

Solution Generation. The decision making scenarios are all about generating solutions to the ethical problems in this case. Use them for starting points.

Testing. If you are doing any formal ethical testing of solutions, make sure to remember that risk is inherent in all forms of treatment. Eliminating risk is not an option. Remember also that if the machines are taken out of service, people will be harmed because of the need for rescheduling, use of other machine or other facilities.

Implementation. Fritz Hager's decision point about notifying other users is a kind of whistle-blowing. You might evaluate both his decision points from the perspective of our discussions on ethical dissent.

•Decision-Making Scenarios.

Decision points 1 and 2 are set inside AECL, and presume knowledge of the historical narrative, but also some technical knowledge of how the errors were generated (see the beginning and advanced technical lessons). Users of decision point two would profit from an earlier exercise that makes them aware of the complexity of the socio-technical system. Instructions for implementing decision points are available in the toolbox.

Decision points 3 and 4 concentrate on difficulties faced by Fritz Hager, the medical physicist at the Tyler TX facility. They presume a knowledge of the historical narrative, but also of the perspective piece on treatment facilities. Less extensive knowledge of the socio-technical system is needed for these decision points than for the first two.

Decision Point 1: Is it fixed?

At this point,[7] the managers asked Don Knott and his team of support engineers for their considered opinion on the safety of the system. Had they made it impossible for the defect to reappear? Had they fixed *the problem*, or simply fixed *a* problem, leaving it

[7] This paragraph is a fictional insertion. This question likely was asked, or it should at least have been asked.

possible for other defects to produce harm in the future? Knott was given 2 days to come up with a plan to answer this question. Much of the future of the Therac-25 depended on a reasonably speedy, accurate answer. His short list of questions begins with:

1) What sort of review would be needed to adequately answer this question? Who would do it?
2) What aspects would be reviewed (software, hardware, design, operations, etc.) and how extensively?
3) What timeline and effort would best balance the need for a thorough review with the need for a reasonably speedy answer?
4) What other steps would he recommend AECL and Therac-25 users take while the plan was being put into action?

Have teams construct plans for Don Knott to present to his management. These need at least to include answers to the four questions posed in the decision point. The more detailed the plans, the more practice students get in moral creativity. This might be done as a homework assignment, as a classroom exercise, or some combination.

Decision Point 2: A Corrective Action Plan.

After the Tyler deaths, AECL was rocked with news of yet another massive overdose at the Yakima, WA facility. After an investigation, AECL identified another software flaw (see "a beginning technical lesson"). After identifying this flaw, the FDA requires that AECL submit what is called a "corrective action plan" or CAP. Listed below are the items that are required in a CAP:

1) Identification of product
2) Number of products affected
3) Notification to user of hazard and instructions for use pending correction
4) Repair, replace or refund of product
5) Reimbursement of expenses to users
6) Specific changes planned to remedy defect
7) Plans to verify that remedy is effective
8) Steps to limit reintroduction of defect into commerce
9) Timetable for completion of actions
10) Proposed dates for progress reports
11) FDA approval of plan

As the chief engineer, Don Knott[8] was a primary participant in drawing up this plan. Some items were easy (e.g. 1) but other required much painstaking detail, along with communication with others outside AECL.

Have teams construct at least an outline for a CAP. Classroom presentations of these will allow other teams to critique and evaluate each other's approaches and learn the variety of ways the problem might be approached.

[8] Again, parts of this paragraph are fictional. It is very likely that Don Knott was involved in designing the CAP, but we have no direct evidence for it.

Decision Point 3: Report the Problem or Deal with it Internally
Immediately upon hearing of the first accident, Hager came back to the ETCC facility to determine what was wrong. Local ETCC management had already been told of the accident, and they had communicated with the umbrella organization that owned ETCC. Both higher and local management were convinced that they could handle the incident locally, without need to report the incident to the Texas Radiological Control Board, AECL, or the FDA. Their main concern was a defensive one: they wanted to avoid making public statements that might be taken as admission of liability or guilt in a lawsuit. Hager knew that there were legal requirements for reporting to the state board, and he also knew that *something* had happened to Cox, though it was unclear what or how severe the injury was. He was convinced that the best thing to do for the safety of patients was to report the incident, both to AECL and to the state. But how could he make a convincing argument for this that overweighed his superiors concern for liability?

A simple classroom brainstorming discussion (with evaluations of the proposed options) of this decision point may be enough to give student some practice generating arguments in favor of Hager's position. But other approaches might involve role play with preparation time to construct an argument, or teams who present arguments to the class for evaluation.

Decision Point 4: Going Public on your own.
Fritz Hager was at ETCC when he got the call that another patient had been hurt by the Therac-25 radiation therapy machine. He rushed over to the treatment room in time to help the parient, Verdon Kidd off the machine and to talk with him about his injury. After getting medical treatment for Mr. Kidd, Hager convened a quick meeting with his assistant and with the operator. They notified management at the center, and suspended treatment with the machine for the day. The operator, who had been involved in two incident now, was quite shaken up, and Hager convinced her to take a break for a while. Then he and his assistant began trying to replicate the error that was now clearly repeatable. Working on a comment from the operator that both incidents had involved her editing the parameters of the treatment after an incorrect initial entry, they were able, after a half hour, to get the error to occur consistently.

Hager's first instinct on finding he could replicate the error was to measure the radiation dose. If it was an overdose it would be critical to know how much of one. So he set up the material to measure the likely dose in a human body, repeated the commands that produced the error, and to his astonishment discovered that this error involved administering 5,000 rads to the patient: a clearly lethal dose. So, the AECL engineer had been wrong after the first incident; Therac-25 *could* administer lethal doses of radiation, and it likely had done so twice in his shop.

After talking with his management to let them know the extent of the problem, he contacted AECL. On the phone, he described to Don Knott, the chief AECL engineer, how to reproduce the error. After a bit he got a call back saying they could not reproduce it. Hager told Knott that timing of the editing was critical, you had to quickly move the

cursor back up to do the edits. With this additional information, Knott cold reproduce the error.

Knott told Hager that AECL would contact the other sites that used Therac-25, warn them of the difficulty, and propose a short term fix. Specifically, he said AECL would:

1) Notify sites that there was a technical problem with the editing functions in the machine,
2) Recommend that all site pry off the up-arrow key, or put tape between its contacts, to render it inoperable,
3) Distribute a more appropriate fix when the AECL engineers constructed one.

Hager expressed his concern that this fix did not tell users about the severity of the problem, or that the problem might result in fatal overdoses. But the three point plan was all he could get Knott to commit to.

Earlier, Hager had imagined his career was over: he had been the medical physicist at a site that had probably lethally overdosed two patients. But having found how to replicate the error, he was at this point pretty sure that it was a design flaw in the machine itself. If it was a design flaw, then any of the other sites using the machine could replicate the error too: and perhaps injure or kill another patient. He had just been at a "user's group" meeting of medical physicists who had charge of Therac-25 machines on their sites. So he had a contact list of all the sites and their personnel. AECL was not going to tell people of the severity of the problem, perhaps he should.

Again, classroom discussion of this decision point may be sufficient. It might concentrate on generating alternative solutions to Hager's difficulty. As you can see from his interview (and the resources section), Hager decided to call other users to make them aware of the problem and to ask them to replicate the error. He did this without checking with local management.

•*Role-Playing Exercises.*

Instructions about setting up role-playing exercises are included in the toolbox. Some role-play you light set up include:

- Don Knott's presentation to his management of his plan, with critique from management
- The conversation between Don Knott and Fritz Hager about what actions AECL would take in notifying users.
- The conversation between Fritz Hager and his management over whether the first accident should be reported.

•*Debating Activities*

Instructions for setting up various kinds of debate are included in the toolbox. Here are several propositions that might be used for debate:

1) The central flaw in the design of Therac-25 was not accounting for race conditions.
2) The central flaw in the design of Therac-25 was removing all hardware safety interlocks in favor of software.
3) Fritz Hager's decision to contact other users to notify them of the error violated his duty to management at the treatment facility.

Hughes Aircraft

Hughes Aircraft Abstract

When computer chips are embedded in expensive weapons systems, the chips need to be tested to make sure they can withstand years of exposure to the extreme environmental hazards they might face (rapid changes in temperature, severe shock, changes in atmospheric pressure, etc.). These chips are sealed in metal containers to protect them from the environmental stress. The seals and the chips need to be tested to make sure they can withstand the stress. Unfortunately, the need to manufacture and deliver these chips on time can compete with the desire to test them thoroughly.

In the mid 1980s, Hughes Microelectronics was manufacturing what were called hybrid microchips for use in guidance systems and other military programs. A series of environmental tests were specified by the government contract. But pressure to ship chips out on time to customers got in the way of complete testing. "Hot" chips, those needed right away for shipment were given preferential treatment by some in charge of the testing process and shipped without the proper tests being performed.

This case is about what happened when employees of Hughes Microelectronics noticed that these tests were being skipped. The decisions they made to report this make this one of the classic cases in the history of whistleblowing.

Historical Narrative

Overview

In the mid 1980s, Hughes Microelectronics was manufacturing what were called hybrid microchips for use in guidance systems and other military programs. A series of environmental tests were specified by the government contract. But pressure to ship chips out on time to customers got in the way of complete testing. "Hot" chips, those needed right away for shipment were given preferential treatment by some in charge of the testing process and shipped without the proper tests being performed.

When computer chips are embedded in expensive weapons systems, the chips need to be tested to make sure they can withstand years of exposure to the extreme environmental hazards they might face (rapid changes in temperature, severe shock, changes in atmospheric pressure, etc.). These chips are sealed in containers to protect them from the environmental stress. The seals and the chips need to be tested to make sure they can withstand the stress. Unfortunately, the need to manufacture and deliver these chips on time can compete with the desire to test them thoroughly.

This case is about what happened when employees of Hughes Microelectronics noticed that these tests were being skipped. The decisions they made to report this makes this one of the classic cases in the history of whistleblowing.

Background

Most of the chips that Hughes Microelectronics was making were of a special sort called "hybrids." Hybrid chips combine two different kinds of semiconductor devices on a common substrate. These hybrid chips are then hermetically sealed in metal or ceramic packages so they are protected from environmental stress and isolated in an inert atmosphere of helium and nitrogen. There were over 70 programs for which Hughes Microelectronics was manufacturing hybrid chips from 1985 to 1987. The chip for each program was different. Because of military secrecy as well as company secrecy, exact specifications of the chips are unavailable. But we provide an example of the sort of chip that was likely among the Hughes chips, an analog to digital converter.

The chips had to be tested not only for whether or not they worked correctly, but for whether or not they held up to standards in terms of their seal or their resistance to heat and shock. The records that Hughes kept regarding their testing showed that approximately 10% of the chips tested failed one or more tests. When a test fails, it does not mean the chip is bad. It might work fine, in fact. But if the seal is broken, water or air might get in over time and corrode the connections on the chip. The tests included things like various programs of temperature cycling, shock tests, and leakage tests.

The next section provides short summaries of five incidents in which Margaret Goodearl and Ruth Ibarra witnessed attempts to to bypass the appropriate tests. More detail is provided for some of the incidents (linked to the incident title).

The Various Incidents

Margaret Goodearl and Ruth Ibarra are the two whistleblower in our case. Goodearl was in charge (along with Donald LaRue) of the floor area in which the testing was done. Ibarra was a quality control agent hired by the company to provide an additional audit of the accuracy and completeness of the tests.

The Lisa Lightner Incident

Lisa Lightner was an operator in environmental testing who conducted leak tests. In August of 1986, Donald LaRue ordered Lightner to pass a hybrid that she had tested to be a "leaker." Lightner, along with Goodearl, reported the incident to upper management. Goodearl was later threatened with loss of her job if she did not reveal "who the squealer was."

The Shirley Reddick Incident

Shirley Reddick was a worker in charge of sealing the lids onto the hybrid packages as well as the stabilization bake process. In October of 1986, Reddick had been ordered by Donald LaRue (a floor manager) to reseal some hybrids. A hybrid is not allowed to be resealed unless it has gone through a complicated and lengthy process, and a "decap" sticker had been placed on it. Reddick complained to Goodearl, who complained to upper management and she was again threatened with loss of her job.

The Rachael Janesch Incident

In the same month (October 1986), LaRue asked Rachel Janesch, another tester in the environmental area, to sign off a leaker as passing the leak test. Goodearl became involved in the reporting of this incident, and the parts were re-tested.

The PLRS Incident

Goodearl and Ibarra found a tote box of PLRS (Position Locating Reporting System) hybrids. PLRS most likely involved some sort of radar function. There was some blank paperwork on the lot travelers accompanying the PLRS parts, meaning that tests had not been run on them before they were passed on. After she reported this incident, Goodearl was told that she was not a part of the team anymore, that LaRue did not trust her, and that her relationship with LaRue was like a divorce in that she was the one that was going to have to go.

Goodearl attempted to file harassment charges in Personnel following the incident. Goodearl was summoned into the office of a middle manager who had been given the harassment documentation by Personnel. He tore up the harassment charge in front of her, flung his glasses at her, and told her that he was going to fire her if she ever went above him to complain again. After this incident, LaRue was removed from his job and taken out of E-1000 in order to avoid further conflict. But his work still involved supervision of testing chips.

The AMRAAM Incident

Two hybrids destined for an air-to-air missile failed the leak test. LaRue placed these chips on his desk with the intention to pass them on without the test during the evening

when Goodearl was not there. By this time, Goodearl and Ibarra were already talking with members of the Office of the Inspector General and were looking for evidence to prove that Hughes Aircraft was intentionally skipping tests. Goodearl and Ibarra photocopied the documentation from the chips showing that they had failed the leak test. They then replaced the chips and their documentation on the desk where LaRue has left them. A few days later they were shipped to a subsidiary of Hughes. They were intercepted by the Department of Defense. The two parts were subsequently tested are were revealed to be leakers.

The decision to blow the whistle

After Goodearl began to report the incident internally to upper management, Goodearl's performance reviews took a sharp drop. Her earlier reviews had been excellent and she had been promoted to her current position because of them. The feedback she was getting from upper management was clear, she had to shut up and get with the team, or lose her job.

Just before the AMRAAM incident, Goodearl and Ibarra, knowing that the series of incidents was likely to continue, placed a telephone call to the Fraud Hotline of the Office of the Inspector General. After several telephone conversations and face to face meetings, they agreed to begin to look for clear evidence of fraud. After the AMRAAM incident, Goodearl was laid off. Ruth Ibarra was transferred to another position that involved loss of most of her responsibility. She later left Hughes.

Court Battles

The Inspector General's office began an investigation in 1989, as soon as they received the clear evidence from the AMRAAM incident.

After Goodearl was laid of by Hughes in 1989, she filed a Wrongful Discharge suit against them. In 1990, Goodearl dropped this suit in favor of what is called a qui tam lawsuit in cooperation with Ruth Ibarra (now married with the last name Aldred). The two whistleblowers claimed in their suit that Hughes was defrauding the Government in its microcircuit testing procedures. Specifically, the civil suit charged Hughes with "knowingly presenting, or causing to be presented, false or fraudulent claims against the United states, or knowingly making, using, or causing to be made or used, a false record or statement to get a false or fraudulent claim allowed or paid by the Government, and for conspiring to defraud the Government by getting a false or fraudulent claim allowed or paid, in violation of the False Claims Act, 31 U.S.C. §§ 3729-32."

The False Claims Act has been around since 1983, and was designed to allow a citizen to sue a U.S. government contractor for making false of fraudulent claims about the quality of the goods or services the contractor has agreed to provide. It allows the citizen to sue "on behalf of" the government (thus the Latin qui tam). The person suing can recover personally up to 25% of whatever damages are eventually assessed. The bulk of the damages go to reimburse the U.S. government.

Goodearl and (now) Aldred filed the civil qui tam suit because they felt the Inspector General's office was too slow in its own investigation. But in 1991, the Department of Defense charged Hughes in criminal court with willfully conspiring to defraud the Government. The charges were defrauding the DoD by "knowingly and deliberately producing hybrids that had not been tested in the manner specified by contract and the pertinent military specifications…and to make false statements, writings and representations on documents in a matter within the jurisdiction of the DoD."

The civil lawsuit was put on hold while the criminal accusations were settled. The criminal trial lasted a month. Hughes' lawyers constantly battered at the credibility of the two main witnesses, Goodearl and Aldred. They claimed that the only fraud that had been committed was the AMRAAM incident, and that all the other incidents were distorted by Goodearl and Aldred, and the Department of Defense. It was a difficult and ugly proceeding, especially for Goodearl and Aldred.

Outcomes

On June 15th, 1992, Hughes was found guilty of conspiring to defraud the government. Donald LaRue, who had also been charged, was found not guilty. Comments by the jury suggest that they felt LaRue had himself been pressured into his actions, and that the company was to blame.

After being found guilty in criminal court, and after an unsuccessful attempt to appeal, Hughes began to negotiate in the civil suit. They agreed to a settlement in 1996. Hughes was assessed 4.05 million for their fraud. Goodearl and Aldred were awarded $891,000 of this amount (22%). Hughes also had to pay the legal fees for Goodearl and Aldred ($450,000).

Both Goodearl and Aldred were profoundly affected by their decision to blow the whistle, and by Hughes retaliation. Goodearl and her husband had to file for bankruptcy, and Aldred was on welfare for a year before she could find another job. Goodearl's marriage eventually broke up. Still, both felt they had been correct in blowing the whistle. After the final settlement, Aldred said, "Despite the toll it has taken, it was the right thing to do."

Time Line

1979	Ruth Ibarra begins working for Hughes Aircraft company's Microelectronic Circuit Division (Hughes MCD) in Newport Beach, CA
1981	Margaret Goodearl begins working for Hughes MCD as a supervisor for assembly on the hybrid production floor and as a supervisor in the hybrid engineering lab
1984	Ibarra becomes supervisor for hybrid quality assurance
1985	Goodearl asks Ibarra to look at errors in paperwork, Ibarra brings errors to the attention of her supervisors and was told to keep quiet, beginning of time period when Goodearl/Ibarra became aware of problems in hybrid chip testing and paperwork
1986	Goodearl becomes supervisor for seals processing in the environmental testing area, False Claims Act (31 U.S.C. §§ 3729-3733) becomes False Claims Reform Act of 1986 making it stronger and easier to apply
Oct. 1986	Goodearl/Ibarra report problems to Hughes management, and, after the problems were not fixed, Goodearl/Ibarra reported the allegations of faulty testing to the United States Department of Defense
Jan. 9, 1987	Earliest date that Hughes may have stopped neglecting environmental screening tests
See Criminal Suit Timeline	
1988	Ibarra leaves Hughes feeling that her job had been stripped of all real responsibility
Mar. 1989	Goodearl is laid off from Hughes
1995	Goodearl and her husband are divorced
See Civil Suit Timeline	
Civil Suit Timeline United States of America, ex rel. Taxpayers Against Fraud, Ruth Aldred (was Ibarra), and Margaret Goodearl v. Hughes Aircraft Company, Inc.	
1990	Goodearl files wrongful discharge suit against Hughes and a number of individual managers, which was eventually dropped in favor of the civil suit
May 29, 1990	Thinking the government investigation was taking too much time, Goodearl/Aldred file civil suit against Hughes under False Claims Reform Act of 1986 with the help of Taxpayers Against Fraud and Washington law firm Phillips & Cohen.
Dec. 1992	Under provisions of the FCA, the U.S. Department of Justice Civil Division takes over the civil case
Sep. 10, 1996	Hughes found guilty in civil trial, to pay U.S. Government $4,050,000 and each relator $891,000 plus a separate payment of $450,000 to cover attorney's fees, costs, and expenses.
Criminal Suit Timeline United States of America v. Hughes Aircraft Co., and Donald LaRue	
Dec. 13, 1991	after a lengthy investigation, the U.S. Department of Defense charges Hughes and Donald A. LaRue with a 51-count indictment accusing it of falsifying tests of microelectronic circuits (criminal suit)

Jun 15, 1992	Hughes found guilty of conspiring to defraud the U.S. Government in criminal case, co-defendant LaRue acquitted following 4-week trial, Goodearl/Aldred called as witnesses in trial, Hughes appeals
Oct. 29, 1992	Hughes fined $3.5 million in criminal trial decision
Dec. 2, 1993	Appellate court upholds 1992 criminal conviction and sentence, Hughes appeal fails

Perspective Pieces

Introduction to Hughes Microelectronics Division

At the time of the incidents we are investigating, Hughes Microelectronics Division was a division of Hughes Aircraft. Hughes Aircraft in turn was owned by General Motors, a major automotive corporation. Hughes Aircraft was originally started in 1932 by the multimillionaire Howard Hughes as a division of Hughes Tool Company. During World War II, Hughes Aircraft became the dominant entity and grew to enormous size as a result of its Defense Department contract to produce radio equipment. After the war, Hughes branched out into radar systems, radar guided missiles, video and infrared imaging, and thermal detection. The company was therefore heavily invested in microelectronics equipment, and began manufacturing its own microelectronics for its systems. Thus was born Hughes Microelectronics. Hughes was one of the original players in the rise of the use of computing technology in defense.

During the 1980's, Hughes Aircraft was one of the top defense contractors in the nation. Hughes Microelectronics was producing chips that were used in at least 73 different military programs during the time from 1985 to 1987. The programs are very important, and lethal systems: F-14, F-15, and B-52 aircraft, guided missiles, radar systems, satellites and tanks. The list covers every branch of the military and many other major defense corporations.

The chips that Hughes Microelectronics was manufacturing were shipped to all these programs as customers. Some "customers" were really other divisions of Hughes Aircraft, and other customers were other defense contractors who were using Hughes parts to produce their own systems for the US government or other purchasers of arms.

Multi-year and multi-system contracts of the kind that Hughes Microelectronics had with the government were worth billions of dollars to Hughes and to its parent companies. So it was clearly in Hughes best interest to meet the guidelines of the contracts.

Some guidelines, however, can exert more immediate pressure than others. Customers (including the US military) call and ask where the chips are that are late, and when they will be delivered. These are immediate questions and need to be responded to immediately.

On the other hand, the outside inspections of the chip manufacturing process are only scheduled at predictable intervals and are announced in advance. Most of the inspection of the process of testing is done internally, by Hughes employees, who report to those who ultimately also have to answer to the customers who are waiting for late chips.

Directors, supervisors, and managers thus have more immediate pressure on them to "deliver the chips" than to make sure that every test on the chips is done. This is the reality of the choices that workers like Margaret Goodearl and Donald LaRue and their supervisors in Hughes Microelectronics Environmental Testing area were facing.

Frank Saia's perspective

Frank Saia has been a long time employee of Hughes Aircraft, and is currently faced with one of the most difficult decisions of his career. He was having problems in the environmental testing phase of his microchip manufacturing plant, and the problems were making him late in delivering the chips to his customers.

Saia began his career at Hughes 35 years ago in 1951. As a physics major from Boston College, he took a job on the east coast for a few years, but soon was enticed to move out to California to work for Hughes. As the electronics age began, Saia was one of the engineers at Hughes whose job grew and changed as the technology he designed and manufactured changed.

Today, he is in charge of manufacturing for the entire microcircuit product line that Hughes produces. This means hundreds of different kinds of microchips, and thousands of versions of those chips on the manufacturing floor at any one time. Several hundred people report to him, indirectly, through what seems to be a slowly increasing group of assistant managers and general supervisors.

Hughes makes computer chips for the US military. And the chips Saia was in charge of making would be used in many different military applications, including F-14 and F-15 fighter aircraft, air-to-air missiles, the M-1 tank, Phoenix missiles, etc. Many of the chips were part of guidance systems for missiles or targeting systems for tanks and aircraft. These battlefield systems undergo tremendous environmental stress from dust, vibration and impact, heat and cold, and long term exposure. Thus the chips needed to be able to withstand these environmental pressures for the life of their service. This is where the environmental testing group came in. They tested the chips before they were sent out to their customers, often other divisions of Hughes who were assembling aircraft or weapons. They, in turn sold the assembled aircraft and weapons to the US government.

Because of the time pressure to deliver chips, Frank Saia had been working to make the production of chips more efficient without losing the quality of the product. Chips are manufactured and then tested, and this provides two places where the process can bottle up. Even though you might have a perfectly fine chip on the floor of the plant, it cannot be shipped without testing. And, since there are several thousand other chips waiting to be tested, it can sit in line for a long time. Saia devised a method that allowed them to put the important chips, the "hot parts," ahead of the others without disrupting the flow and without losing the chips in the shuffle. This let hot parts get through faster and meant they could meet the order volume they needed.

But Saia was not only concerned with getting parts through quickly. When a subordinate suggested they cut a test he had added, his reply was " It is the worst thing you can do to ship bad parts." The test he had added both helped to assure quality parts and to make the testing go more quickly. It was called the "gross leak" test and it could quickly tell if a chip in a sealed container was actually sealed or not. Adding this test early in the testing sequence allowed them to not waste time testing chips that would fail a more fine grained leak test later in the sequence.

Saia was proud of his reputation as a problem solver. He had another reputation too, one that often worked in his favor, but of which he was less proud. He had a temper. And when the line backed up and parts would be delayed, he made sure everyone knew exactly how he felt about it. Hughes was a military contractor, and they hired many military people. Saia ran his section with military strictness and made sure people did their jobs.

So, when he heard that the environmental testing area was behind again, he called in Don LaRue to let him know how he felt about it. How did he feel? He was angry and he was insistent that he would not be embarrassed by late shipments. Saia was getting regular calls from Karl Reismueller, the director of the Division of Microelectronics at Hughes. Reismueller made it clear that the parts had to get out the door. In addition, Reismueller had given Saia's telephone number to several of the customers for the chips, whose own production lines were shut down awaiting the parts that Saia was having trouble delivering. His customers were now calling him directly to say "we're dying out here" for need of parts.

Don LaRue, the general supervisor in charge of the environmental testing area, was sure to be unhappy any time that Frank Saia was unhappy. They both began to look for ways to speed up the delivery of chips. LaRue was already putting "hot parts" at the front of the line for testing, and this was not enough. Saia applied more pressure. He told LaRue to baby-sit the parts all the way through the process, from one test to the next, and make sure they pass.

But now Saia has heard that LaRue has actually been skipping tests. Since LaRue began this practice, they have certainly been more on time in their shipments. Besides, both LaRue and Saia knew that many of the "hot" parts were actually for systems that were in the testing phase, rather than for ones that were being put into active use. So testing chips for long-term durability that went into these systems didn't matter. But still, LaRue had been caught by Quality Control skipping a test, and now he needed to make a decision. Upper management had simply told him to "handle it" and to keep the parts on time.

Decision Point

He couldn't let LaRue continue skipping tests, or at least he couldn't let this skipping go unsupervised. LaRue was a good employee, but he didn't have the science background to know which tests would do the least damage if they were skipped. He could work with LaRue and help him figure out the best tests to skip so the least harm was done. But getting directly involved in the skipping of tests would mean he was violating company policy and, likely federal law. His bosses seemed to have little patience for his explanations that the environmental testing took more time than the manufacturing. They did not believe that his hard work has made the line as efficient as it could be. They wanted results (which meant chips out the door) now. So, he had to keep the pressure on LaRue to get the chips out the door. But he did feel like he had one choice to make. He could keep the pressure up and simply turn a blind eye to LaRue's practice of test skipping. Or, he could use his expertise to match the test skipping with the particular chip and its application so the most timesavings were achieved and the least risk was incurred.

Life on the Testing Line

Most of the "girls" on the testing line had a high school education, and some had previous experience in precision manufacturing. But the work of testing chips was really something you learned right there on the job. You learned the job from another girl mostly, though you had some supervision from the floor managers (LaRue and Goodearl).

There were about 14 girls on the line, and though they were collectively called "the girls" a few were male. Each girl was assigned a station she learned, and those who had been there for some time knew the work on several stations. The work on each station consisted of doing the testing and keeping records of the fact that the testing had been done. Each chip traveled about in a pink plastic bag, and each bag had a "lot traveler" attached to it. This traveler specified what tests were to be done, and what the particular settings needed to be on that test. The chips traveled about the floor in plastic tote boxes, about 20 at a time, each with its own traveler. There could be different testing values indicated on the travelers for chips in the same tote, so testers had to carefully read the lot travelers to determine the appropriate tests and settings. Thus, each tester:

1. Selected a chip (and its traveler) from the line of chips to be tested at that station. This meant making a decision about which chips were more important, that is, were there any "hot parts."
2. Checked the lot traveler that was attached to the chip.
3. Made sure the setting on the test machine was done as per the specifications on the traveler.
4. Put the chip in the testing device and ran the test.
5. Read the results of the test and marked this on the traveler that went with the chip.
6. Moved the chip on to the next appropriate testing station, or put it in the "rework" bin, or marked it to be discarded (all depending on the type of chip and the outcome of the type of test).

This work was relentless, painstaking, and done under great time pressure. Chips needed to go out to customers, and Hughes Microelectronics (in the person of Don LaRue) would not tolerate any slow or inaccurate work. It was made more difficult by the fact that, at any one time, there were between 500 and 1,000 chips in the testing room. Each of these chips had its own testing regimen outlined on its traveler. But even chips that were the same might be marked differently on the traveler, since in addition to "production parts" that were shipped to customers, some were simply for the folks in engineering to use or for proof of design (does it work) or manufacture (can we make it). These chips required less testing.

Matters were not made better by the management style that was standard at Hughes. People would be told to do things, given no reason, and were expected to jump to the task immediately without question. If you did not, you might be warned once, or, or occasion, simply fired without warning.

Quality assurance (QA) was a regular presence in the area, though staff for this was thin. The QA people would look for mistakes in procedure and report these to supervisors. On occasion, the US government would make (carefully announced) visits to audit the testing area to make sure everything was being done in accordance with procedure.

But the most regular presence in testing was that of the floor manager Don LaRue. In addition to overseeing the testing room, his primary job was to make sure "hot parts" got tested quickly and got shipped to the customers on time. When his assistant, Margaret Goodearl arrived, he was able to spend more time babysitting the hot parts. If a chip failed a test, he would often take the chip from a tester and retest it himself, sometime after hours. He regularly pushed "hot parts" through the testing procedure and explained little of what he was doing to the girls. He would simply take a part from them and disappear. He had work to do, and was simply too busy to explain.

Margaret Goodearl

Margaret Goodearl was born in Ireland and finished high school in Ireland and England before she emigrated to the US. In California, she worked doing assembly and then line management for several precision manufacturing companies (including a company that made mechanical heart valves). When she got her job at Hughes Microelectronics Division, she was first employed in manufacturing, but was moved from that job because she could not get the security clearances she needed. She was not yet a US citizen.

She was eventually promoted to a supervisory position in the environmental testing group. She would be working with Donald LaRue, the current supervisor for environmental testing. LaRue was near retirement and would need someone to take over his position when he was gone. She was promoted to be his assistant with the idea that she would take his position upon his retirement. The group that LaRue (and now Goodearl) supervised tested the chips that Hughes made in order to make sure they would survive under the drastic environmental conditions they would be likely to face.

Hughes made computer chips for the US military. The chips Goodearl was in charge of testing would be used in many different military applications, including F-14 and F-15 fighter aircraft, air-to-air missiles, the M-1 tank, Phoenix missiles, etc. Many of the chips were part of guidance systems for missiles or targeting systems for tanks and aircraft. These battlefield systems undergo tremendous environmental stress from dust, vibration and impact, heat and cold, and long term exposure. Thus the chips needed to be able to withstand these environmental pressures. The chips were protected against the environment by a metal cap that would seal off everything about the chip except the connectors from the harsh environment. But the sealing process itself could damage the chips, or might not have been done correctly, or the chip might fail because of some internal flaw that would become evident only under stress.

This is where Margaret Goodearl's testing group came in. They tested the chips after they were sealed and before they were sent out to their customers. Often other divisions of Hughes were the customers and were assembling aircraft or weapons from the parts they

received. These customers, in turn, sold the assembled aircraft and weapons to the US government.

In an ideal world, the chips Hughes was sending to its customers (and thereby on to the government) would be subjected to rigorous environmental tests. Those chips that failed would either be reworked (if the contract allowed it) or scrapped. Only thoroughly tested chips would be delivered, because of the critical nature of the military systems involved.

At first, things were fine. She shared a desk with LaRue on the floor of the testing area, with a view of all the testing stations. She followed LaRue around the environmental testing area, learned how to do all the tests, and became acquainted with all the "girls" that she would be supervising (they were called this even though a few were male). She learned how and when to make exceptions to the required tests, and learned all the different protocols associated with each test. She learned how to quickly read a "lot traveler," or the paperwork that accompanied a chip as it was being tested. The traveler specified what tests the chip required and had a place for the test operator to check of that the chip had passed.

There was one additional wrinkle in this standard procedure: hot parts. Some of the chips going through the testing procedure were behind schedule and were in short supply. Customers were regularly calling upper management with complaints that these parts were not being shipped quickly enough. So, Don LaRue met every morning with representatives of upper management to determine which part were "hot" that day and needed to be rushed through. He would then make sure that these parts were tested first and shipped along quickly, ahead of chips that could wait.

Decision Point

So the ideal was rigorous testing of the chips, but with some chips getting in line ahead of others to be tested. In the last several months, however, things had not been ideal. Goodearl saw tests being skipped. When she tried to report these problems, she was told she was not a team player, that she should simply do her job and pass the chips, and that if she kept reporting problems, she might get fired. But because the tests were so important, and the chips were for such crucial systems, she felt like she had no other choice but to continue to report the problems she saw. At that point, it began to get really bad.

Goodearl and the Lisa Lightner Incident

A few months after Margaret Goodearl started her new position, she was presented with a difficult problem: one of the girls, Lisa Lightner, came to her desk crying. She was in tears and trembling because Donald LaRue had forcefully insisted that she pass a chip that she was sure had failed the test she was running.

Lightner ran the hermeticity test on the chips. The chips were enclosed in a metal container, and one of the questions was whether the seal to that container leaked or not. From her test, she was sure that the chip was a "leaker" -- the seal was not airtight and water and corrosion could seep in over time and damage the chip. She came to Goodearl for advice. Should she do what LaRue said and pass a chip she knew was a leaker?

Goodearl suggested they take the chip together to the Quality Assurance people and tell them the story. Quality Assurance (QA) was the group whose job was to oversee the manufacture and testing of the chips. They had the authority to make this decision. Goodearl knew one of the people in QA, Ruth Ibarra. After consulting with Ibarra, Goodearl and Lightner decided to keep the chip for the moment and make an appointment with Karl Reismueller, the head of the entire Division of Microelectronics.

They were not successful, and were told they would need to go through channels. They did get a meeting with Richard Himmel, the manager of the Microelectronics Circuit Product Line. Lightner was afraid for her job, because LaRue had told her she would lose it if she disobeyed him. Himmel assured her that her job was safe, and told her she should try to work with LaRue. Still, he said, she was not required to pass parts that she knew did not pass the tests.

Lightner was somewhat calmed by this conversation, but Goodearl's life got worse rapidly. Goodearl got a call less than an hour later from a very upset manager, Frank Saia. Saia was the direct supervisor for LaRue, and thus also Goodearl's supervisor. Saia was known for his temper and displays of anger, and after letting her know how unhappy he was, he demanded to know "who the damn squealer was out there. If I don't hear from you by 4 o'clock on this, you're fired." Just before she took that call, she has walked by Saia's office and seen LaRue leaving it, crying. Clearly things were bad.

Later that afternoon, she received a phone call from Jim Temple, an assistant manager under Saia. Temple reminded her of her immigrant status, and said that meant that if she was fired, she would likely end up cleaning toilets for a living.

Decision Point

Later that week, there was a large meeting of all the "girls" and management. The testers were told that they needed to obey LaRue, but that no one would be asked to do anything against the rules. Clearly, to management, the incident was over. The original chip, by the way, was given back to Don LaRue, who passed it.

Ruth Ibarra and the Role of Quality Assurance

Ruth Ibarra was the supervisor of quality inspection in the environmental testing area. She began working in quality assurance (QA) in May of 1984. QA was responsible for certifying that appropriate procedures were followed in the testing of the chips. Because she was constantly in the environmental testing area, Ibarra knew Margaret Goodearl and her supervisor, Donald LaRue well. It was natural, then, for Ibarra and Goodearl to talk with each other regularly about how testing was proceeding.

Every hybrid chip had to come through QA for initial screening tests (e.g., is the paperwork correct? does it correspond to the chip to which it is attached?) before being sealed. If any rework (for instance, resealing the chips) was to be done on the hybrids, they had to pass through QA again for further screening tests where the circuitry of the hybrid was compared to the circuitry of an illustrated model. Finally, QA gave the

paperwork of each part the final inspection to verify that they had all been done properly before the hybrids were sent to the customer.

Ruth Ibarra's job in QA was to watch over the proceedings in the environmental testing area. She supervised the QA people who did the initial checks and who did the final checks on the paperwork before the chips left environmental testing. In the process of walking around in among her workers, Ibarra's job was to check on how the tests were being followed in environmental testing. She did not supervise the "girls" in environmental testing, but she was there as an inspector. The main tool she used in her inspections was a close reading of the paperwork that followed each chip as it went through the testing process.

Each chip took about 10 days to get through the entire process of testing. Paperwork called a "lot traveler" traveled through the process with each chip. The lot traveler specified what kind of chip it was, and what tests it should undergo. These lot travelers were the center piece of the whole process. Everything that was to be done to the chips was specified in the lot traveler, and once it was done, needed to be noted on the lot traveler. When the part left the factory, the lot traveler stayed behind as the authoritative record of what had happened to that particular chip. Thus, falsifying a lot traveler was like lying about what tests were being done.

The role of Ibarra and other supervisors was made more difficult by the fact that some chips that were quite similar might have completely different testing routines. For instance, about 2% of the chips being tested were called "proof of design" chips that engineers were working on. These would not be shipped to customers, but were for internal use as the engineers tried out different designs. The engineers occasionally wanted these new designs tested, and so would send the chips down to testing with their own lot travelers specifying the tests they wanted. These would be tested differently, and perhaps with more loose standards than the chips that were being sent to customers.

To make things more difficult still, every day some parts in the testing room would be labeled "hot" by the management. This meant these needed to be rushed through the process ahead of other chips so they did not spend the usual 10 days getting tested. This allowed Hughes to rush the chips out to customers who needed them quickly. Donald LaRue was in charge of seeing that the hot chips got special treatment and went ahead of other chips in the testing process. The chips were still supposed to be given the tests required by the lot traveler, but if they were "in line" for a test with chips that were not "hot" they would go to the front of the line, and sometimes be hand carried from test to test so they could be shipped quickly.

This pressure between doing the tests correctly and getting parts out in a hurry ended up causing tension between Quality Assurance and Environmental Testing. Specifically, it caused tension between Ibarra, in charge of Quality Assurance, and LaRue, in charge of seeing that "hot chips" got shipped as soon as possible. Margaret Goodearl was caught in this tension and sided with Ibarra, in favor of slowing the line down so there were no

mistakes in testing. This internal conflict of speed versus quality was the center of the disputes that eventually arose.

Ibarra, Goodearl, and the Shirley Reddick Incident

Several weeks after the Lisa Lightner incident, Ruth Ibarra (from QA) was walking through the environmental testing area and saw Shirley Reddick resealing some chips. This was a normal thing to do, but all chips for reseal need to have a stamp on them indicating this was authorized. None of these chips had the stamp. Ibarra, suspicious because of the things she had heard of the Lightner incident, decided to ask Goodearl what was going on. Goodearl did not know, so they went back to Reddick, who was still doing the resealing, to ask who authorized the resealing. "Don LaRue did," said Reddick, but she did not know why. Goodearl then asked Don LaRue why the resealing was being done without the stickers, and received the reply "None of your damn business." At this point, Goodearl decided to stop her questioning.

But things were not over yet. Later that day, she received a phone call from Jim Temple, one of her superiors, telling here to come to his office. Temple informed her in no uncertain terms that she needed to back down. "You are doing it again. You are not part of the team, running to Quality with every little problem." Goodearl insisted she did not "run to Quality" but that Quality came to her with the concern. Temple was unmoved. "Shape up and be part of the team if you want your job."

At this point, Goodearl decided to talk with the personnel office to inquire about making a harassment complaint regarding the threats of firing. After her meeting there, she saw the person immediately walk down the hall to Frank Saia's office, the head of the entire product line on which she worked. She then got a call to come to Frank Saia's office. She had had a run-in with Saia in the earlier incident, so as she went to the meeting she was nervous. Saia ask her to sit down, and then erupted, throwing his glasses across the room, in her direction: "If you ever do anything like that again, I will fire your ass right out of here."

Later that week, at a company dinner meeting, she spoke with the head of the personnel department, Mr Neiendam, who assured her that her job was not at stake and that she did not have to worry about Saia or LaRue. After another week passed, she heard from LaRue himself that he had been transferred to another department, Production Control, but that he will still be moving chips in and out of the environmental testing area.

The chips that started this incident, by the way, were resealed and sent on to the customer.

Goodearl's new boss, B.J. Rhodes, was actually assigned the position in addition to her other duties, and so Goodearl was mostly left alone in the Environmental Testing area, with regular visits from Don LaRue, who in his new role in Production Control was still requiring the girls to skip tests. When Goodearl reported this to her new boss, B.J., her boss said, "That's none of your business. Your job now is to turn the people around on the

floor and make them like you. Upper management doesn't want to hear about this. Don't make any more waves, you've made enough problems. Just do your job."

As the production control person, LaRue continued to give "hot parts" to the girls and get the parts special treatment in passing tests.

Decision Point

Goodearl, Ibarra, and the AMRAAM Incident

Now that Goodearl had few sympathizers among upper management, she increasingly turned to Ruth Ibarra in Quality assurance for support in her concerns about test skipping and the falsification of paperwork.

One day, Goodearl noticed that some AMRAAM chips with leak stickers were left on her project desk in the environmental testing area. The leak stickers meant that the seal on the chips' supposedly airtight enclosure had failed a test to see if they leaked. AMRAAM meant that the chips were destined to be a part of an Advanced Medium Range Air-to-Air Missile. Goodearl knew that these parts could not be retested and needed to be simply thrown away. So why was someone keeping them? She also knew that these were officially "hot parts" and that the company was behind schedule in shipping these parts.

After consulting with Ruth Ibarra, the two of them decided to do some sleuthing. They took the chips and their lot travelers to a photocopy machine and made copies of the travelers with "failed" noted on the leak test. They then replaced the chips and their travelers on the desk. Later that day, as Don LaRue passed the desk, Goodearl asked Don LaRue if he knew anything about the chips. "None of your business," he replied. The chips disappeared, and later the travelers showed up in company files with the "failed" altered to "passed." So, Goodearl and Ibarra had clear evidence (in their photocopy of the "failed" on the traveler) that someone was passing off failed chips to their customers. And these were important chips, part of the guidance system of an air-to-air missile.

Decision Point

Supporting Documents

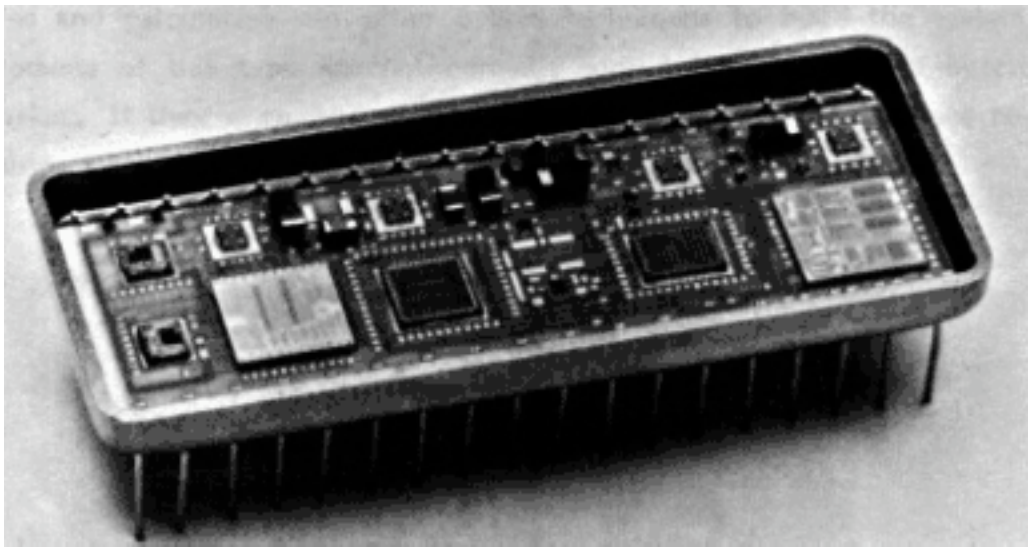
Historical Documents

Look for a photocopy of a lot traveler.

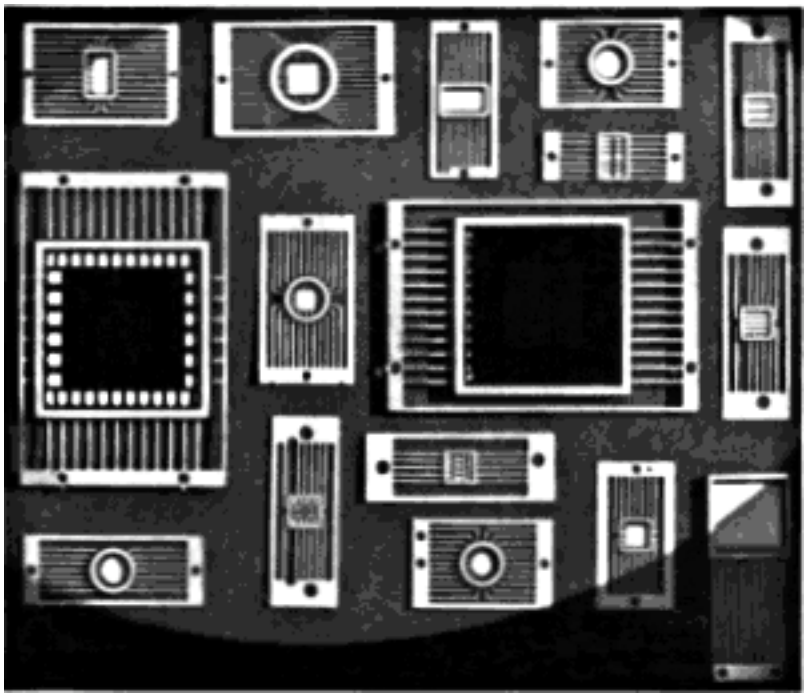
Maps, Tables, & Figures

Pictures

•An unsealed hybrid microcircuit.



•Examples of Hybrid Packages from the Mid-80's



•Select Positions in Hughes Organizational Chart

Select Positions in Hughes Organizational Chart	

Karl Reismueller Division Manager	
Olaf Neiendam Head of Personnel Department	Richard Himmel Manager of Microelectronics Circuit Product Line
	Frank Saia Manager of Microcircuit Manufacturing
	Jim Temple Assistant Manager of Hybrid Production
	Donald LaRue General Supervisor: Seal Symbol and Phase Shifter
Ruth Ibarra Quality Assurance	Margaret Goodearl Assistant General Supervisor of Environmental Area
	Lisa Lightner Shirley Reddick and an additional 16-20 Test Operators

Other resources

Hybrid microelectronics at Hughes

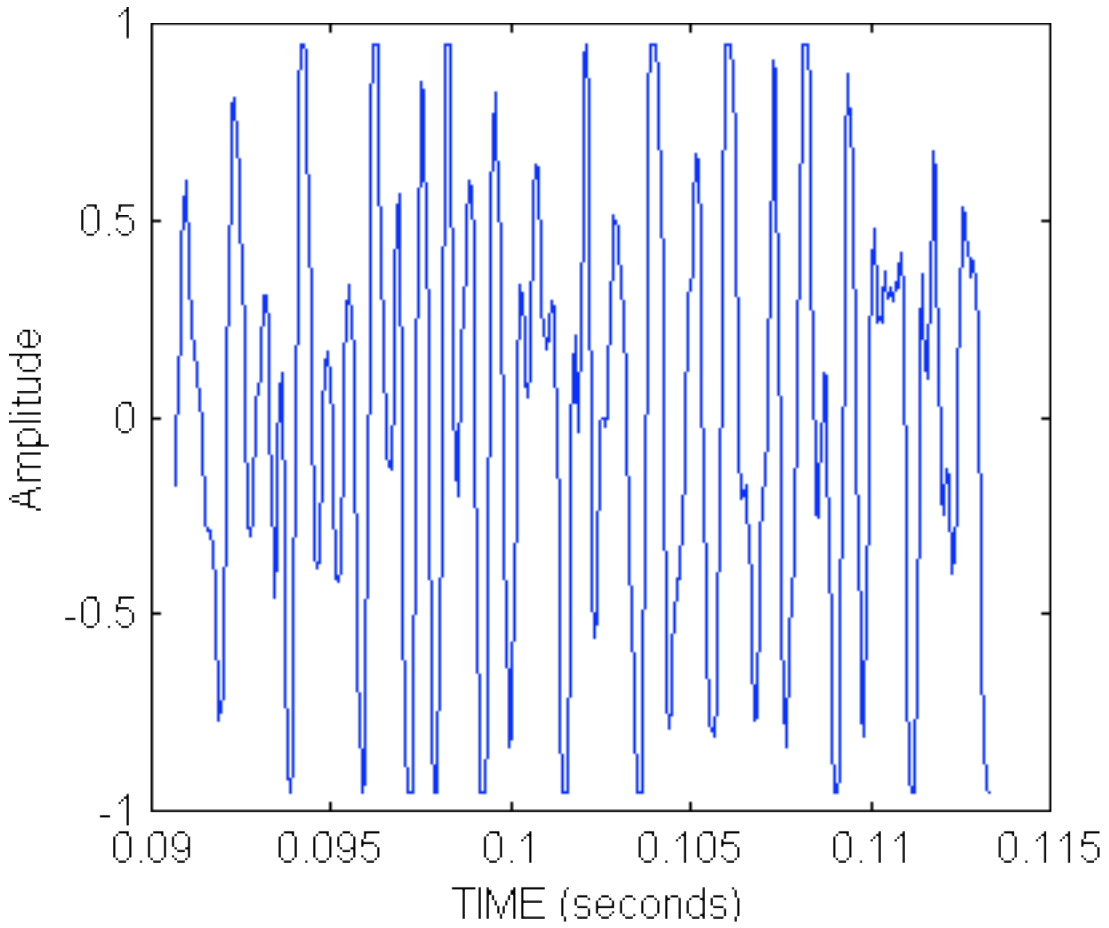
Most of the chips that Hughes Microelectronics was making were of a special sort called "hybrids." Hybrid chips combine two different kinds of semiconductor devices on a common substrate. For use in the high-stress world of weaponry, these hybrid chips are then hermetically sealed in metal or ceramic packages so they are protected from environmental stress. The circuitry is thereby isolated in an inert gas atmosphere of helium and nitrogen. This sealing protects the chips from corrosion and other environmental damages.

There were over 70 programs for which Hughes Microelectronics was manufacturing hybrid chips from 1985 to 1987. The chip for each program was different. Because of military secrecy as well as company secrecy, exact specifications of the chips are unavailable. To get some idea of what the chip Hughes manufactured did, we present here an Analog-to-Digital (A/D) converter. This was a common hybrid that was among those Hughes was manufacturing at the time.

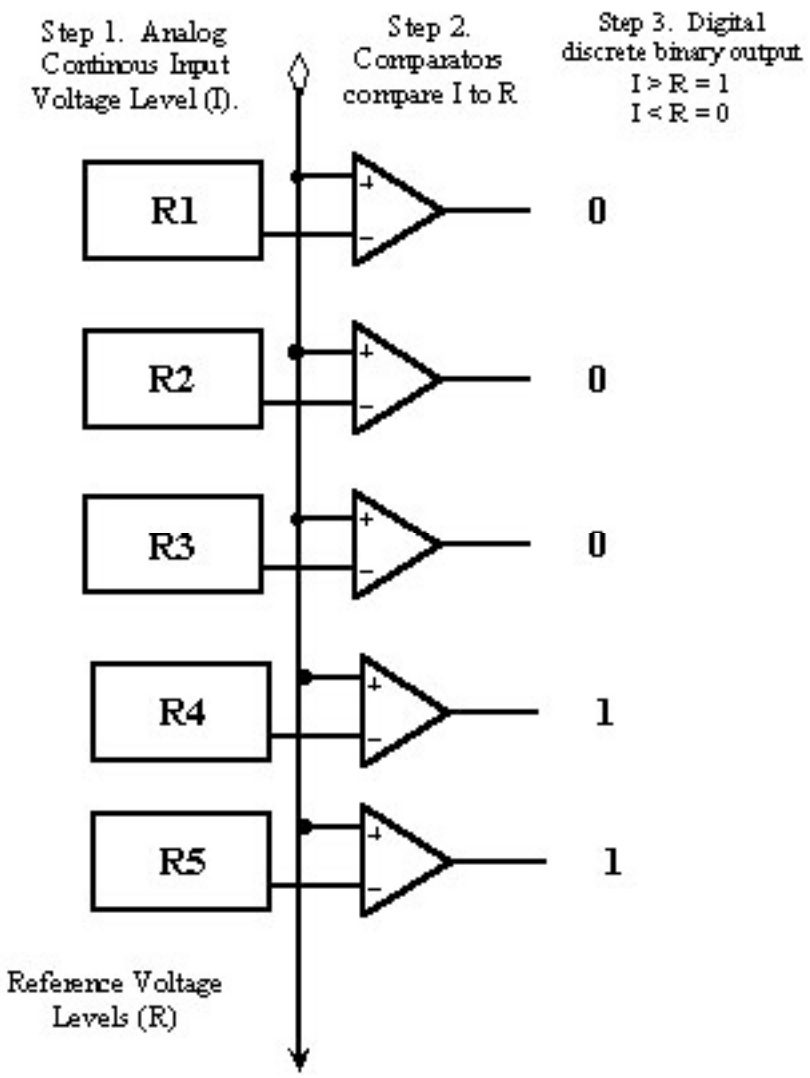
•Analog-to-Digital Conversion basics

One purpose of Analog-to-Digital conversion (A/D conversion) is to change a continuously variable analog signal into discrete digital signals that can serve as input to a computer.

An analog signal is a continuously variable physical signal. It can take many forms: radio wave, cellular phone transmission, radar signal etc. Usually, in A/D conversion the analog signal takes the form of an electrical current. An electrical current has a continuously changing voltage, and, when illustrated, usually takes on a sinusoidal wave form. The complex sine wave you see here is a combination of several waves, some of which might be "true" signal and other noise.



A digital signal is a representation of this wave form as a discrete set of numerical values. All data that can be understood by central processing units (CPU) are in this discrete binary form. To understand how hybrid A/D converters function, it is important to note how A/D conversion works on the circuits. The circuit receives an analog signal (usually an amount of voltage) and simultaneously compares it to a set of reference voltages.

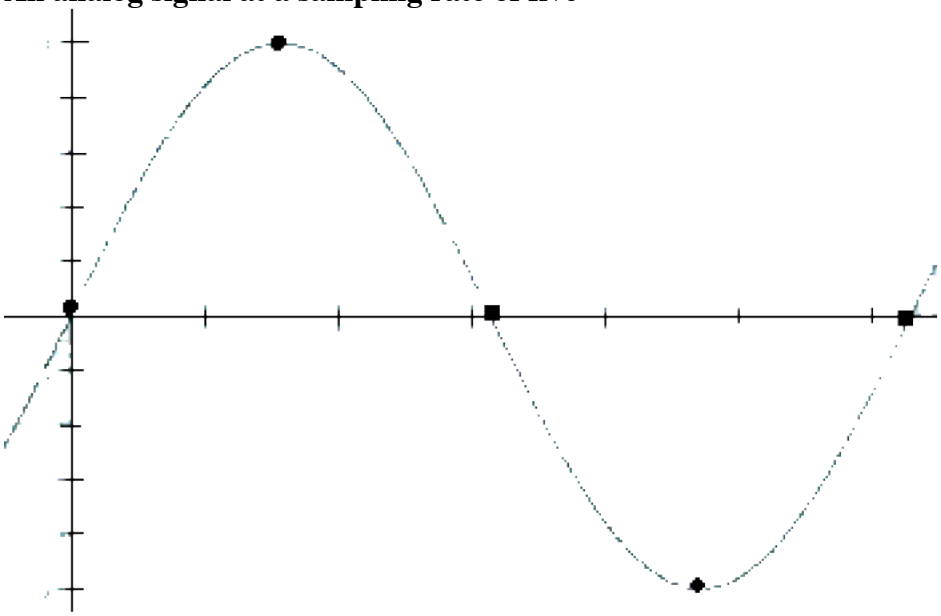


The number of comparators vary by the speed of the hybrid microcircuit (8-bit, 12-bit, 16-bit etc.). Each comparator is one Least Significant Bit (LSB) higher than the comparator immediately below it. If the input voltage level is above the reference voltage, the comparator takes on a "1" value. If the input voltage is below the reference voltage, then the comparator remains a "0" value. The output is thus in binary code as a sequence of zeros and ones that show where the voltage was at that instant in time. The process described above occurs at a specific sampling rate measured in Megahertz (megahertz is a million cycles of electromagnetic currency alternation per second). The input voltage level continues to change and is regularly sampled a certain amount of times per second. These discrete samples are then run through the comparators which produce the binary code.

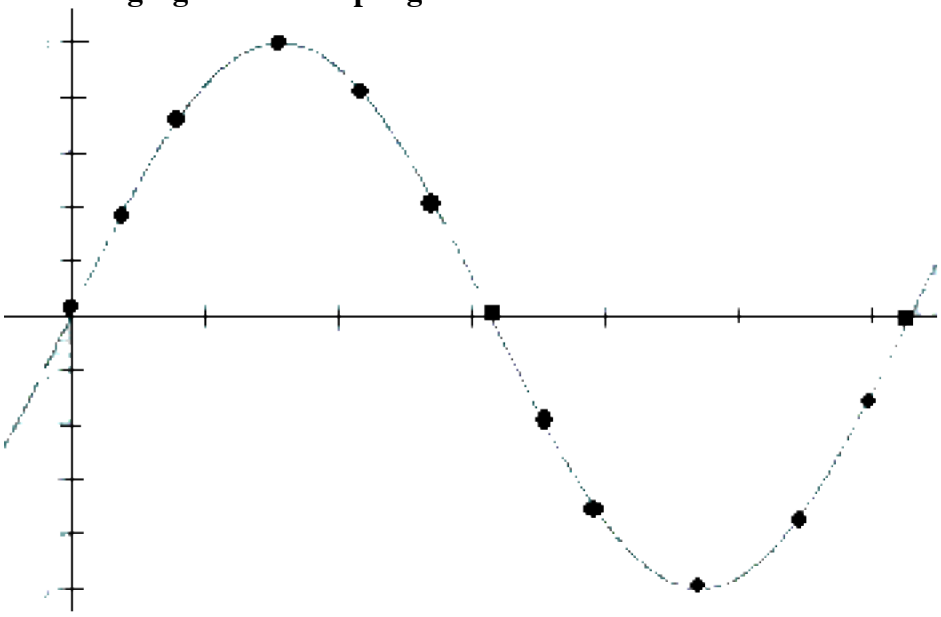
When the binary code switches from 0 to 1, the actual voltage level is somewhere between the fixed levels of the two comparators. So, the more comparators there are in the circuit, the smaller the differences between the fixed voltage levels in the comparators

can be. The closer the comparators fixed levels are, the better the computer's guess is to the original input voltage level.

An analog signal at a sampling rate of five



An analog signal at a sampling rate of thirteen



In addition, the more frequently the samples are taken, the closer the digitized representation of the voltage wave is to its actual form. You can see this easily by thinking of sampling occurring every time there is a dot in this figure vs. every other time there is a dot. The faster sampling rate gives a much better picture of the incoming signal.

Increasing the density of sampling and increasing the density of comparators are the two factors that can make A/D conversion more accurate.

So, a hybrid A/D chip has two kinds of technology on it. There has to be technology to receive, amplify, and relay analog signals. There also has to be technology to sample the signals and run the samples through comparators in order to change them to binary code.

These complex chips were then sealed in metal or plastic (see pictures in the resource section), so that the extreme variation in battlefield environment would not damage them. If they leaked when they were shipped from the factory, they could become damaged more easily in the extremes of battlefield environment. And if they became damaged (e.g. by heat, or cycling from heat to cold, or by shock), they could fail in a variety of ways. They might give a wrong signal or they might give no signal at all. If they gave a wrong signal, this might not be detected, and the missile might be mis-targeted or the airplane guidance system might give incorrect readings. In the heat of battle, or even in training runs, this could have lethal consequences.

Programs affected by Hughes

The hybrid microcircuits manufactured by Hughes aircraft co. were used in a large variety of programs affecting every branch of the military and other cutting-edge technological government programs. The legal documents from the cases filed against Hughes list 73 programs for which Hughes manufactured hybrid microcircuits between 1985 and 1987 (see the list below). They do not state that every program was affected by the fraud, but simply that these were the program that used hybrids chips, and that hybrid chips were fraudulently passed off as having been appropriately tested.

If they are marked with an asterisk (*), they are one of the 27 programs that were definitely affected by Hughes during the relevant period of time. The other programs could have been affected, but no clear evidence was obtained to prove this. The lack of evidence may mean they were not effected, but it may also mean documents were falsified to show that some programs were not effected.

Not all parts of this chart can be filled in at this time for various reasons. We have filled in those we can.

Hughes Acronym	Item	Client
*F-14	Fighter jet (Tomcat)	NAVY
·*M-1	Tank	ARMY
·*SVS	Missile Defense	Boeing
·R5		
·*A6E	Grumman Intruder	
·*TPQ-37	Firefire weapon locating system	Army (Raytheon)
·REG	Range and Approach Guidance	
·*PHXMIS	Phoenix Missile	Navy
·*DESC	Energy Supplier	Defense Energy Support Center

·BLUESKY		
·*LORS3A	Loral S3A	Loral Space and Communications
·M5T1		
·*LHD	Amphibious Assault Ship	NAVY
·*ASO F-14	Aviation Supply Office F-14	Navy
·*AMRAAM	Advanced Medium Range Air-to-Air Missile	AIRFORCE
·*SADS	Submarine Antenna Distribution System (SADS) or SADS (Simulated Air Defense System)	
·ADCAP	Torpedoes	Navy
·*LORF-4	Loral F-4	Loral Space and Communications
·*PLRS	Position Locating Reporting System	Dept of Transportation
·*TPQ-36	Firefinder Radar	Marine Corps
·J4TA		
·*F/A-18	Fighter jet	NAVY
·*ELE 6		
·*AHEP		
·*MULE	Modular Universal Laser Equipment (MULE)	Army
·*AMRPWR		
·*GLLD	Ground Laser Locator Design	
·N2		
·*HADR	Hughes Air Defense Radar	
·*F-15	Fighter jet	AIRFORCE
·*F-15 MSIP	Fighter jet Multi-stage improvement program	AIRFORCE
·TEXINST	Military semiconductors	Texas Instruments
·LASERHAWK	Laser augmented Airborne TOW (Tube-launched, optically tracked, Wire-guided	ARMY
·SPD	Power supplies	SPD
·ASO NAVY	Air Support Operations	NAVY
·DIVADS	Gun System	Division Air Defense
·OTHER 45		
·LITAPPTE		
·OTHER 14		
·MVS		
·ROBINS		

·MAVOFF		
·*SPERRY	Designed and built computers for military and aerospace	Sperry Rand's
·*JTIDS	Joint Tactical Information Distribution System	
·VENUS MAP		
·INTEL SAT	International TeleCommunications Satellite	
·PRKELN		
·ASO A6E		
·GOES	Geosynchrous Orbital Environmental Satellite	
·NORD	Unmanned Military Aerospace Vehicles	Nord
·ASARS2	Advanced Synthetic Aperture Radar System	
·IPD/TAS	Improved point defense/target acquisition system	
·KUBAND	Antenna used on the space shuttle	NASA
·OTHER 94		
·B-52	Bomber jet	AIRFORCE
·GMS	Gun Management System	
·LTD	Laser Target Designator	
·MCDOUG	Thor Missile	McDonnell Douglas/Boeing
·PHXMOD		
·*RAYTHN	Merged with Hughes Aircraft	Raytheon
·ASO F-18		
·GMDELCO		Hughes Delco Operations
·EATAS		
·HYBRID		
·SILCONIX	Semiconductors	Vishay Siliconix
·K8		
·T1		Airforce
·HS250		
·R6548PS		
·OTHER 15		
·HRBSIN		
·B52 CP	B52 Command Post	

A quick glance over the names of the programs and companies or branches of the military effected gives you an idea of how important these hybrid microcircuits are. Missiles, fighter jets, a bomber, radar systems, and an assault ship are extremely lethal and important systems. If something malfunctioned because of a faulty hybrid, the consequences could be dire indeed.

To give a more specific example of this, we describe the AMRAAM (#15) in terms of its history, capabilities, relation to the hybrid microcircuit, and the potential effect a faulty chip might have on it.

•The AMRAAM and potential effects of chip failure

The Advanced Medium Range Air-to-Air Missile (AMRAAM) development program started in 1975. [28]. In February 1979, the AMRAAM program completed its conceptual phase. The United States Air Force (USAF) selected two companies as competing contractors to continue to develop the AMRAAM – Hughes Aircraft Co. and Raytheon Co. (Raytheon). Thirty-three months later in December 1981, both companies successfully demonstrated the effectiveness of their prototypes.

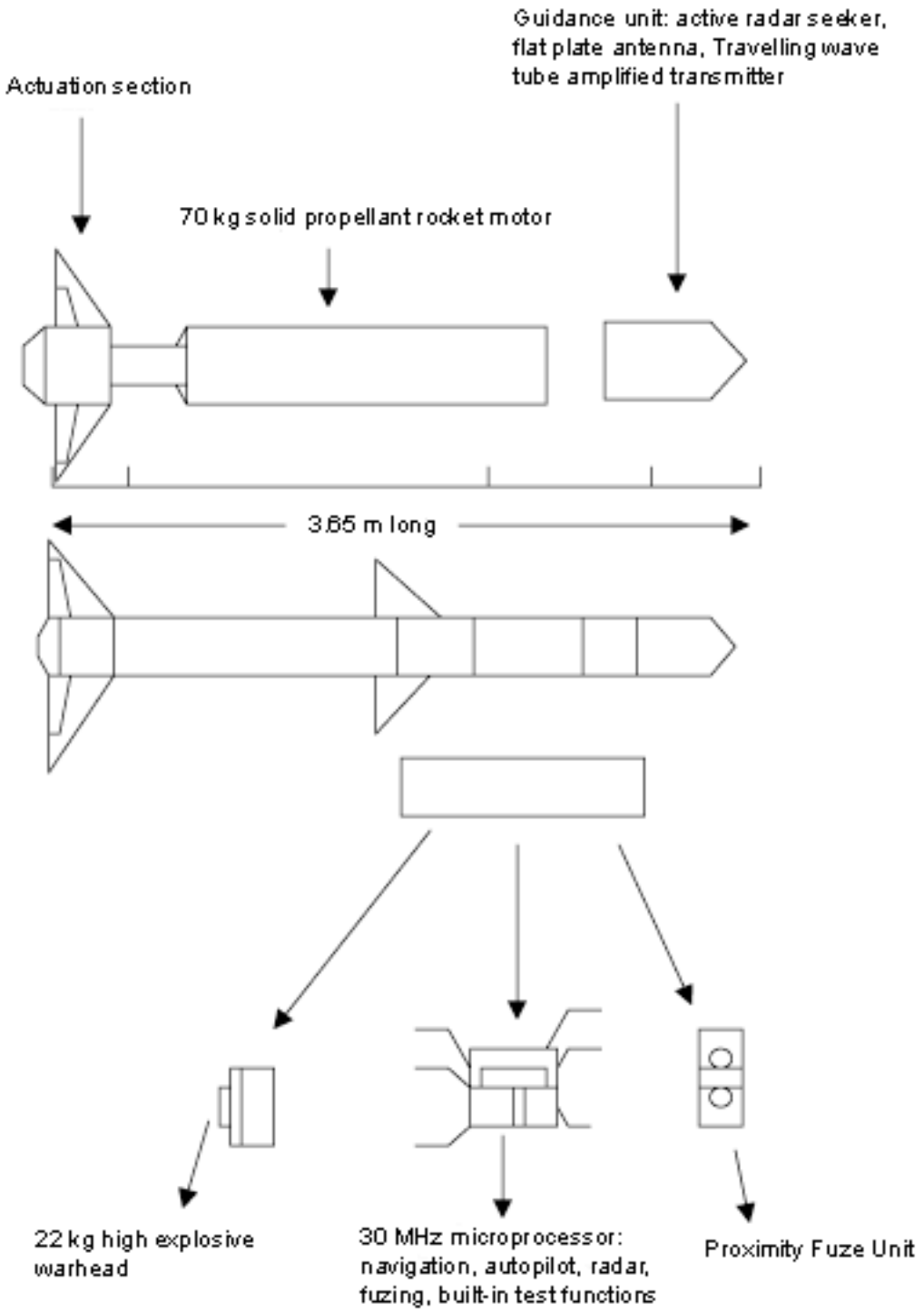
The AMRAAM was designed to outpace its predecessor the AIM-7 Sparrow by having higher speed, greater range, increased maneuverability, better resistance to electronic counter measures, an active terminal radar seeker, and improved reliability and maintainability. [28]. American was selected as the full-scale developer. Test missiles were launched sometime after full-scale development was complete, and kills in Operation Southern Watch and Bosnia proved the AMRAAM's capability. [29]. The AMRAAM entered service in 1991 after it was delayed by development problems.

The AMRAAM is a lethal missile that can be launched from the USAF's F-15 (Figure 1), the Navy's F-14, Germany's F-4, Britain's Sea Harrier and other aircraft from over seas Allied forces.

3.65 meters in length and weighing 157 kg before launch, the AMRAAM carries a 22-kg high explosive hollow charge blast effect warhead. The blast effect warhead does not explode upon impact with its target, instead when the missile senses it is within lethal range, it self-detonates. Even more, it is a directed fragmentation warhead filled with 198 separate rod-shaped projectiles. "It is reported that the proximity…system can sense which side of the missile the target is on and direct the blast/projectiles towards the target rather than being distributed in an even, circular pattern. [28]. The AMRAAM can travel at super-sonic speeds, and it can fly, in some cases, up to 40 miles to its target. [29]

Most importantly though, the AMRAAM has a built in radar system so the pilot of the F-15 or F-14 does not need to be an active participant in its guidance. Instead, they can concentrate on more important maneuvers like evading enemy fire. After leaving the range of the plane's guiding radar, and once it is within range of the enemy plane it goes into autonomous mode. This self-guidance system works using a technique called Semi-Active Radar Homing (SARH). Pulses of radar signal are sent out of the missile's head instead of a continuous stream of radar so the target cannot lock on to the missile's signal and administer counter measures. The missile's seeking system follows the target designated by a radar-lock from the warplane, and then it follows its own radar. This system virtually assures a kill when working properly.

The structure of the AMRAAM can be seen here:



Semi-Active Radar Homing and other functions of the missile's microprocessor did not evolve without the important microchip development of Hughes in the early 80's. Hughes was a pioneer in developing a hybrid microchip that was most likely used in the AMRAAM's 30 MHz microprocessor. Hughes manufactured these microchips, which are

actually contained in airtight, hermetically sealed containers, for all branches of the military in the 1980's. Hughes was supposed to test the containers according to the standards the Department of Defense required. Among other tests detailed in the standards, the box was to be tested for correct operation following exposure to extremes in temperature, vibration, pressure, and electrical shock. It was important that these tests be run, because one error in a decimal place in analog-to-digital conversion meant possible disaster.

The AMRAAM's missile head contains hybrid microchips. Using self-guided, Semi-Active Radar Homing, the AMRAAM receives analog radar signals that have bounced off the enemy plane. The microcomputer uses a digital code translation of the analog radar signal to plot the trajectory of the enemy plane. The microcomputer then sends the digital signal to a digital-to-analog converter chip. The information sent by the microcomputer tells the missile where it needs to be in order to intercept and destroy the plane. The new analog signal is then involved in moving the AMRAAM's rudders to correct it's own trajectory to intercept that of the enemy plane. This process occurs every few milliseconds and millions of times over the course of the missile's flight.

In light of the Hughes fraud, it is important to ask what could happen if the hybrid chips were not functioning correctly and the United States went to war. A pilot relying on an AMRAAM to destroy an enemy plane would be at a major disadvantage if his/her missiles were slightly off-calibration. In addition, there is the problem about where the missile goes if it does not hit the target. AMRAAM's have built in self-destruct capabilities if they are not locked on to a target. They have a maximum range of 50 km and a minimum range of 2 km. It would be rather important to feel safe firing a 22 kg high explosive warhead missile at an enemy plane 2km away from you. The speeds that these planes are traveling can make up that distance in seconds. Unforeseen explosions are a huge potential distraction for not only the plane that fired the missile, but other planes in the area.

These musings about potential failure of guidance systems can be replicated for each of the 27 programs we know were affected by the fraud, and perhaps for all 73 programs. The AMRAAM sounds more catastrophic because the guidance systems are working in vehicles traveling at very high speeds. But there are likely lethal consequences even for an Amphibious Assault Ship whose guidance systems go awry.

Although Hughes was convicted of fraud in not testing the chips, there is no way we can be sure the chips were actually faulty (or actually fine), short of finding all the chips in all the weapons systems listed above and checking each one. This is obviously prohibitively expensive, so we still do not know how many (if any) leaking or otherwise badly compromised.

Testing the Chips

•Why test?

The chips that Hughes was manufacturing for the U.S. government were being used in a range of military programs some of which involved aircraft and missiles that traveled at high altitudes at supersonic speeds. Chips that were in these aircraft or missiles were exposed to dramatic ranges of temperature, moisture, and physical shock. So, a chip that worked fine now might not work fine after two years of exposure to the sort of abuse that regular flight exposed it to. The seal might begin to break, or solder links might crack, and the chip would begin to malfunction.

So, the chips had to be tested not only for whether or not they worked correctly, but for whether or not they held up to standard in terms of their seal or their resistance to heat and shock. The records that Hughes kept regarding their testing showed that approximately 10% of the chips tested failed one or more tests. When a test fails, it does not mean the chip is bad. It might work fine, in fact. But if the seal is broken, water or air might get in over time and corrode the connections on the chip.

As mentioned before, the two whistleblowers noticed that Hughes was not fulfilling its military contract in testing its hybrid microcircuits. The two whistleblowers worked as supervisors of hybrid quality assurance and seals processing in the environmental testing area respectively. They were present when Hughes omitted tests, ran tests out of order, authorized re-work on chips that had failed tests, and falsified documents thereby covering up the fraud.

•The Tests

The 4 test areas that were specifically mentioned in case documents as places where Hughes fell short of their contract were Temperature Cycle, Constant Acceleration or Mechanical Shock, Hermeticity (Fine-Leak and Gross-Leak Tests) and the P.I.N.D. Test.

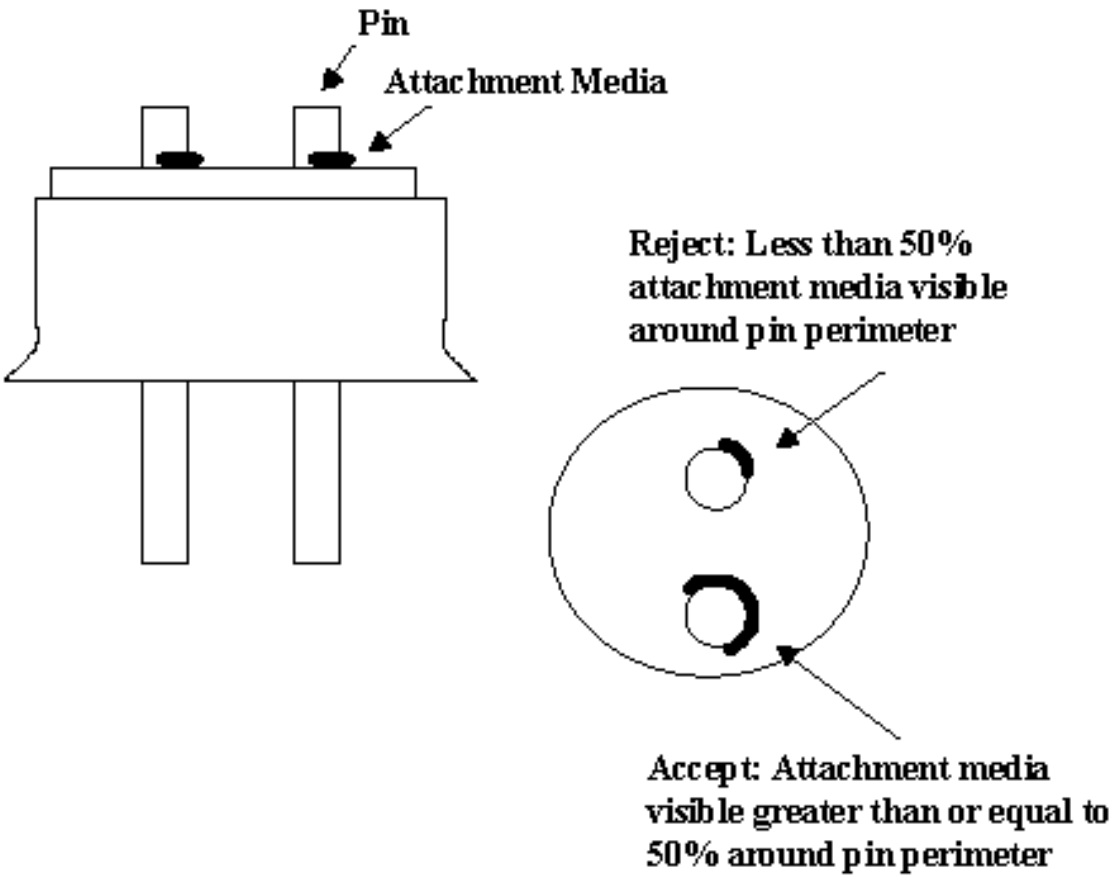
These test descriptions are taken from MIL-STD-883, a more recent version of the test standards than was used during Hughes manufacture in the mid-80s. Although it is possible that some of the testing standards have changed some since Hughes was prosecuted for faulty procedures, all of the basic tests described in the civil complaint are still required today. MIL-STD-883 and can be downloaded in its entirety (641 pages) in PDF format at this web address: http://www.dscc.dla.mil/Programs/MilSpec/listdocs.asp?BasicDoc=MIL-STD-883 or at http://www.dscc.dla.mil/ and link to the Military standards page. Search for MIL-STD-883 on the search engine linked at the bottom of the first Military Standards page.

•Precap Visual Inspection: Method 2017

"The purpose of this test is to visually inspect the internal materials, construction, and workmanship of hybrid, multichip and multichip module microcircuits. This test will normally be used on microelectronic devices prior to capping or encapsulation on a 100 percent inspection basis to detect and eliminate devices with internal defects that could

lead to device failure in normal application." The diagrams from this method detail many errors that could have occurred during the microcircuit construction process: bonding, wiring, adhesive, configuration and other problems. One example is Figure 2017-3c Package Post Criteria. As you can see, the top pin is not attached properly because the attachment media is not visibly surrounding 50% or more of the pin's perimeter. [27].

Figure 2017-3C



•*Stabilization Bake: Method 1008*

"The purpose of this test is to determine the effect on microelectronic devices of storage at elevated temperatures without electrical stress applied." There are nine minimum temperatures and time conditions that hybrid microcircuits can be subjected to. They are all equivalent to one another, but vary in length and intensity. Three examples of valid conditions are 100 degrees C for 1,000 hours, 160 degrees C for 16 hours, or 200 degrees C for 6 hours. End-point measurements are then run on the hybrids after they have been removed from the heating apparatus within 96 hours.

•Temperature Cycle: Method 1010

"This test is conducted to determine the resistance of a part to extremes of high and low temperatures, and to the effect of alternate exposures to these extremes." The hybrids are cycled between two extremes in temperature. One cycle starts when a hybrid is subjected to —65 degree C temperature between 10 and 15 minutes. Then the hybrid is transferred in less than a minute to a temperature of +150 degrees C for 10 to 15 minutes. The transfer time remains less than a minute when it completes the hot section and is returned to the cold section. This one cycle is repeated a minimum of 10 times.

If the number of interruptions (failure in machinery, failure to transfer the hybrid between cycles in less than a minute, etc.) exceed 10% of the total number of cycles run, the test must be completely restarted. "Failure of end-point measurements, evidence or damage to the case, leads, seals or illegible markings shall be considered a failure."

•Constant Acceleration: Method 2001

"This test is used to determine the effects of constant acceleration on microelectronic devices. It is an acceleration test designed to indicate types of structural and mechanical weakness not necessarily detected in shock and vibration tests. It may be used as a high stress test to determine the mechanical limits of the package, internal metallization and lead system, die or substrate attachment, and other elements of the microelectronic device."

The hybrid is oriented respectively at X1, X2, Y1, Y2, Z1 and Z2 (as if on a 3-dimensional plane). For each orientation, it is spun around in a centrifuge machine for 1 minute at 30,000 gravity units (g's).

•Mechanical Shock: Method 2002

"The shock test is intended to determine the suitability of the devices for use in electronic equipment which may be subjected to moderately severe shocks as a result of suddenly applied forces or abrupt changes in motion produced by rough handling, transportation, or field operation. Shocks of this type may disturb operating characteristics of cause damage similar to that resulting from excessive vibration, particularly if the shock pulses are repetitive."

The hybrid subjected to 5 shock pulses at a 1,500 g level lasting .5 ms each. This procedure is repeated in the X1, X2, Y1, Y2, Z1, and Z2 orientations. "After subjection to the test, failure of any specified measurements or examination, evidence of defects or damage to the case, leads, or seals, or illegible markings shall be considered a failure.

•P.I.N.D. Test: Method 2020

"The purpose of this test is to detect loose particles inside a device cavity. The test provides a nondestructive means of identifying those devices containing particles of sufficient mass that, upon impact with the case, excite the transducer."

The hybrid is subjected to series of 4 alternating 1,000+ or — 200-g peak shocks and 20 g peak at 40 to 250 Hz vibrations. Visual indication of high frequency spikes, audio indication of clicks, pops, and rattling, and more complicated electronic measures on a noise detector which exceed the normal background white noise level indicate a failure. Rejects shall not be re-tested.

•Hermeticity: Method 1014

In the Hughes case, these were referred to as the Fine- and Gross-Leak Tests. In MIL-STD-883E it is referred to as the Seal Test. Whatever its name, "the purpose of this test is to determine the effectiveness (hermeticity) of the seal of microelectronic and semiconductor devices with designed internal cavities."

The hybrids are put through a perfluorocarbon gross leak test, penetrant dye gross leak test, or weight gain gross leak test simply to find out if they are not in an airtight container. The fine-leak tests, which follow one or all of these gross leak tests, are some variation on a tracer gas (He) fine leak test. Examples of failure criteria for the tests involve a leak rate that is too fast, or weight gain, bubbles escaping from the seal, or evidence of dye penetration into the seal.

"Devices which fail gross leak may be re-tested destructively. [This means the device can be broken open, resealed, and then retested -computingcases editor] If the retest shows a device to pass, that was originally thought to be a failure, then the device need not be counted as a failure…Devices which fail fine leak test conditions shall not be re-tested for acceptance unless specifically permitted by the applicable acquisition document."

•Pre Burn-In Electrical

This is just a test in which the hybrid is tested to see if it functions the way it is supposed to function under normal conditions following all of the previous tests. It is not submitted to any adverse conditions.

•Burn-In: Method 101

"The burn-in test is performed for the purpose of screening or eliminating marginal devices, those with inherent defects or defects resulting from manufacturing aberrations which cause time and stress dependent failures. It is the intent of this screen to stress microcircuits at or above maximum rated operation conditions or to apply equivalent screening conditions which will reveal time and stress dependant failure modes with equal or greater sensitivity."

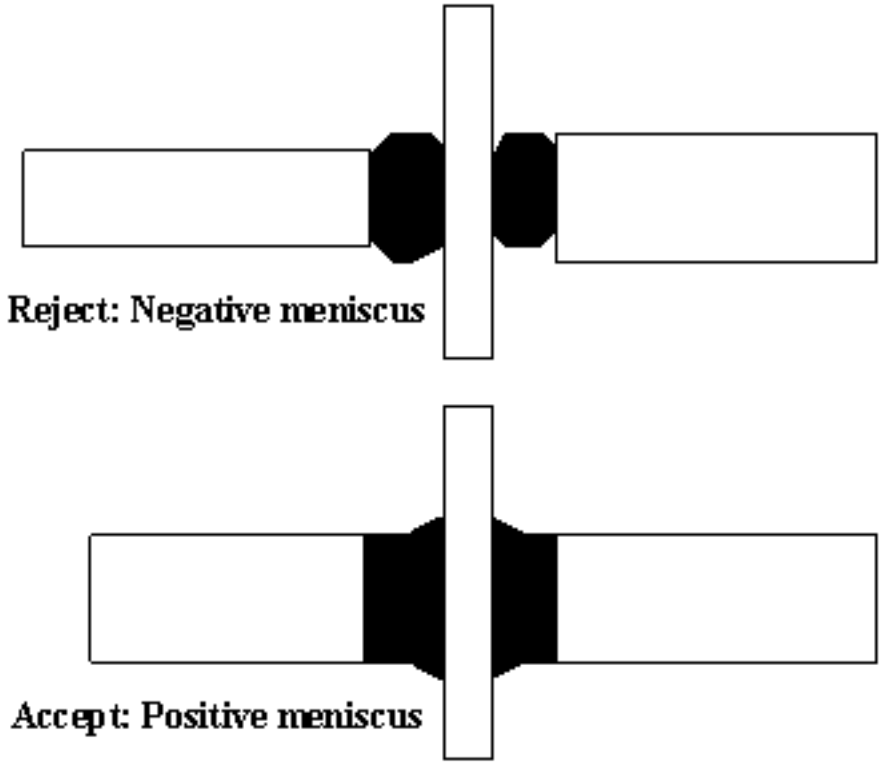
•Final Electrical Test

This test is the same as the Pre-Burn-In Electrical test. The hybrid is tested to see if it functions the way it is supposed to function under normal conditions following all of the previous tests. It is not submitted to any adverse conditions.

•Final Visual Inspection: Method 2009

"The purpose of this test method is to verify the workmanship of hermetically packaged devices. This test method shall also be utilized to inspect for damage due to handling, assembly, and/or test of the packaged device." This test is performed on hybrids when they are about to leave the factory, or have just entered another factory as a finished product. Much like the Precap Visual inspection, this method mainly details specific visual details that result in accepting or rejecting a hybrid. One example is Figure 2009.9 Reentrant Seals. As you can see, only seals with a positive meniscus are not rejected.

Figure 2009-9



U.S. Whistleblower Law

•Protection for Public and Private Employees

•False Claims Reform Act of 1986

(Title 31 — Money and Finance, §§ 3729 — 3733)
In 1863, the False Claims Act was written to provide a civil penalty "of double the amount of damages suffered by the government, plus a $2,000 forfeiture for each false claim submitted." [9]. The law was "enacted to prosecute Civil War manufacturers who substituted sawdust for gunpowder in Union army supplies." [7]

Any person could submit a lawsuit on behalf of the government regarding a false claim against the government. These people are referred to as the qui tam. Qui tam comes from the Latin "qui tam pro domino rege quam pro sic ipso in hoc parte sequitur," meaning "who as well for the king as for himself sues in this matter." Black's Law Dictionary (1979) defines a qui tam action as "an action brought by an informer, under a statute which establishes a penalty for the commission or omission of a certain act, and provides that the same shall be recoverable in a civil action, part of the penalty to go to any person who will bring such action and the remainder to the state or some other institution." [9] In other words, the qui tam plaintiff is sues on behalf of his/her own right as well as that of the government.

Amendments in both 1943 and 1986 were enacted to "increase detection and prosecution of false claims submitted to the federal government." [8] The Reform Act of 1986 was "the brain-child of public-interest attorney John R. Phillips." [7] If the Attorney General elects to take over the case, whistleblowers are guaranteed 15 to 25 percent of funds recovered as well as legitimate compensation for legal fees, back pay, and other damages. If the Attorney General does not elect to take over the case, the whistleblowers are guaranteed 25 to 30 percent of the winnings.

In general, a claim is defined in § 3729 (1986) as, "any request or demand, whether under a contract or otherwise, for money or property which is made to a contractor, grantee, or other recipient if the United States Government provides any portion of the money or property which is requested or demanded, or if the Government will reimburse such contractor, grantee, or other recipient for any portion of the money or property which is requested or demanded."

The Act defines 7 acts that can be prosecuted as false claims. Individuals can be prosecuted under the False Claims Reform Act of 1986 only if they knowingly defraud the government with one of these false claims. By "knowingly", the Act states that a person, with respect to pertinent information about the false claim, "has actual knowledge of the information, acts in deliberate ignorance of the truth or falsity of the information, or acts in reckless disregard of the truth or falsity of the information." This means that violators can be prosecuted not on "clear and convincing evidence" which was required in the 1863 Act but only on a "preponderance of evidence." An employee does not need to prove that their employer submitte4d a false claim, just have a "good-faith belief that a violation had been committed."

False claims as defined by the 1986 Reform Act are as follows:
"§ 3729 False Claims
(a) Liability for Certain Acts — Any person who —

1. knowingly presents, or causes to be presented, to an officer or employee of the United States Government or a member of the Armed Forces of the Unites States a false or fraudulent claim for payment or approval;
2. knowingly makes, uses, or causes to be made or used, a false record or statement to get a false or fraudulent claim paid or approved by the Government;

3. conspires to defraud the Government by getting a false or fraudulent claim allowed or paid;
4. has possession, custody, or control of the property or money used, or to be used, by the Government and, intending to defraud the Government or willfully to conceal the property, delivers, or causes to be delivered, less property than the amount for which the person receives a certificate or receipt;
5. authorized to make or deliver a document certifying receipt of property used, or to be used, by the Government and, intending to defraud the Government, makes or delivers the receipt without completely knowing that the information on the receipt is true;
6. knowingly buys, or receives as a pledge of an obligation or debt, public property from an officer or employee of the Government, or a member of the Armed Forces, who lawfully may not sell or pledge the property; or
7. knowingly makes, uses, or causes to be made or used, a false record or statement to conceal, avoid, or decrease an obligation to pay or transmit money or property to the Government."

The penalties for violations can be very costly. Violators can pay up to $10,000 for each false claim as well as attorney's fees and other costs. But, if the violator admits to submitting a false claim within 30 days of the Government discovering it, the fines are reduced to no less than twice that suffered by the Government. All other damages are waived.

Most importantly, the 1986 revision includes a whistleblower protection provision. (31 U.S.C. § 3730 (h)) "Any employee who is discharged, demoted, suspended, threatened, harassed, or in any other manner discriminated against in the terms and conditions of employment by his or her employer because of lawful acts done by the employee on behalf of the employee or others in furtherance of an action under this section, including investigation for, initiation of, testimony for, or assistance in an action filed or to be filed under this section, shall be entitled to all relief necessary to make the employee whole." [4]

•Protection for Government Employees

•Whistleblower Protection Act of 1989

(Title 5 — Government Organization and Employees, § 1201)
In order to prevent retaliation against whistleblowers, the Civil Service Reform Act of 1978 established the Office of Special Counsel. Since whistleblowers "serve the public interest by assisting in the elimination of fraud, waste, abuse and unnecessary Government expenditures," the Whistleblower Protection Act of 1989 was written in order to strengthen this protection for whistleblowers by the Office of Special Counsel. Congress found that "protecting employees who disclose Government illegality, waste, abuse, and corruption is a major step toward a more effective civil service." The Act improves protection as follows:

1. mandates "that employees should not suffer adverse consequences as a result of prohibited personnel practices and
2. establishes that the primary role of the Office of the Special Counsel is to:
    a. Protect employees who seek assistance
    b. Act in the interest of these employees; and
    c. While disciplining those who commit prohibited personnel practices, remember that protection of employees who seek assistance remains the paramount consideration."

The Act provides deadlines to which the Office of Special Counsel must adhere in prosecuting whistleblower complaints. "To help prevent retaliation against whistleblowers while their cases are pending, the Counsel is specifically prohibited from disclosing the identity of whistleblowers, except when necessary to prevent imminent danger to the public or to prevent criminal activity." [8] To prevent delays after trial, whistleblowers who win their cases are compensated for attorney's fees and other costs while appellate court reviews are pending. [10]

•Protection for Employees of Defense Contractors

•Department of Defense Authorization Act of 1987

(Title 10 — Armed Forces, § 2409)
Much like the whistleblower protection clause in the 1986 False Claims Reform Act, the Department of Defense Authorization Act of 1987 was written specifically to prohibit retaliation against whistleblowers. But, this Act was written specifically for employees of defense contractors who disclose "substantial violations of the law."

Under this Act, "an employee of a contractor may not be discharged, demoted, or otherwise discriminated against as a reprisal for disclosing to a Member of congress or an authorized official of an agency or the Department of Justice information relating to a substantial violation of law related to a contract."

The maximum penalty for violation is complete compensation for damages to the employee. This can included rehiring, back pay, employment benefits, attorneys' fees, or other fees that were lost or otherwise reasonably incurred by the whistleblower throughout the course of "bringing the complaint regarding the reprisal [to the] head of the agency." [11]

How the Hotline for Reporting Fraud Works

The Office of Inspector General for the United States Military has as one of its goals overseeing the manufacture and procurement of military hardware so that the U.S. Government gets its money's worth in these transactions.

The Inspector General runs a hotline for reporting waste, fraud, and abuse. When the problems at Hughes surfaced, Margaret Goodearl had the option of calling a telephone number to report those problems to an outside auditor like the Inspector General's

Hotline. Going outside of an organization's channels like this is called whistleblowing, and though it looks easy to do, it can produce significant hardship for the whistleblower. Supervisors feel like their trust has been betrayed, co-workers can avoid or gang up against you, and the organization can isolate and eventually, fire you.

For this reason, hotline for reporting problems are usually confidential. This is the way the Inspector General's hotline is set up. You can read all about this at the hotline website the Inspector General has set up. Of course, in 1985, when the problems at Hughes surfaced, the web was not an option, but Goodearl and others working at the plant knew about the hotline because brochures about it were passed out.

Here is how the hotline worked then (and now, except you can now use email):

1. The person who wants to lodge a complaint makes the call and is connected to an investigator
2. The investigator will ask a series of questions (who, what, where, when, how, and why) to get the background needed to investigate the complaint.
3. The caller can remain anonymous (no identifying information is given) or can give information so he or she can be contacted for additional information. If contact information is given, it is kept confidential (not shared with anyone)
4. The investigator will assign the case a number for reference and will begin the process of checking on it.
5. In some cases, this results in charges against the offending party, in others, there is not enough evidence to substantiate a case.

If the case is substantiated, it may be resolved internally, or it may be resolved in the federal court system. For more information on how whistleblowing works, see chapter 5.

Analysis Documents

Socio-technical System

Overview

The Hughes case is less about technical computing systems and more about people and procedures in a corporate environment. Thus, hardware and software will get little attention in this analysis. We will spend most of our time thinking about the people and the procedures involved in this case. Even the data structures we will analyze are not those in the chips themselves, but those that help to track the environmental testing of the chips.

•Hardware

Hybrid microelectronics are at the heart of this case. But it is not the chips themselves that are important to the case. Instead, it is the high-stress environment in which the chips would be used: United States military battlefield systems. The chips in air-to-air missiles or F15 aircraft or tanks would be subject to extreme variations in temperature, shock, and other hazards. This is why testing the chips in their sealed containers was so important, and why skipping those tests, or shipping chips that had failed tests, was such a serious fraud.

•Software

The chips themselves embodied software routines, but the design of the software is not the center of this case. Instead, the case is about the organizational procedures that allowed or encouraged the fraud, and that resulted in the whistleblowing.

•Physical Surroundings

The major physical surroundings for this case involve E-1000 at Hughes Microelectronics in Newport Beach CA. E-1000 was a very large, "clean" room containing all the testing equipment, environmental testers, quality assurance, and engineers responsible for testing thousands of hybrids every month. It was an open factory floor style organization, and was designed this way to allow the supervisors to better view all the operations going on. This physical organization enhanced the power of the supervisors, since the operators were always "on display." The section on life on the testing line explains more about the social dynamics of the testing floor.

•People

The various players in this case, from Goodearl and Ibarra to upper management at Hughes to the Defense Department, help us understand the complexity of the case. It is easy to simply look for "bad guys" and "good guys" in this case, but a closer look not only makes the picture more complicated, but helps us think about how to change the socio-technical system in ways that will make fraud less likely and whistleblowing less catastrophic.

Here we attempt to unpack the positions, perspectives, and motivations of the different players in this case. This approach is somewhat like what philosophers call stakeholder analysis. But stakeholder analysis includes a party because they have an interest in a decision that someone might make. Socio-technical analysis includes a party because they are part of a system that influences the actions involved. Thus, stakeholder analysis would be unlikely to include the Personnel Department of Hughes when analyzing Goodearl's decision to blow the whistle. They are not directly affected by her decision. But we must do so here, because Personnel was one of the main players in convincing Goodearl that she had no real options for internal complaint.

•The whistleblowers

Margaret Goodearl and Ruth Ibarra are the main actors in this case, and their decision to blow the whistle on Hughes' fraud is the reason the case is of interest. Goodearl was a newly-appointed supervisor of the testing room for hybrid microchips at Hughes. Ibarra was an employee in Quality control. They describe their motives as ones of concern for military personnel whose lives will depend on the chips they are testing. It became clear in the criminal trial that there was a good deal of tension between Goodearl and her co-supervisor, Donald LaRue, whom she accused of committing the fraud. LaRue's defense made good mileage of this tension, and regularly implied that Goodearl and Ibarra were "out to get" LaRue. In part, the bad relationship may have been a motivation, but as Goodearl and Ibarra tell the story, the bad relationship began when they started reporting that LaRue was skipping tests.

In addition, the structure of a qui tam lawsuit (see US Whistleblower Law) meant that Goodearl and Ibarra were likely to gain a great deal if they won the lawsuit. These lawsuits allow a person to sue a government contractor on the behalf of the US government and to keep some of the money if the suit is successful. In their defense in the criminal trial, Hughes attempted to discredit them by regularly referring to the monetary gain that Goodearl and Ibarra might receive. If Hughes was found guilty in the criminal trial, the civil qui tam lawsuit was likely to go against Hughes also -- and in favor of Goodearl and Ibarra.

Still, there seems little reason to believe that Goodearl and Ibarra were merely "gold-diggers." There is clear evidence that both Goodearl and Ibarra were worried about their jobs. And rightly so, since they were eventually both forced to leave -- Goodearl upon being fired and Ibarra upon being moved to a position with no responsibility. They tried several times to resolve the matter internally. And they paid a great price over the ten years it took to finally reach a settlement.

•Upper management of Hughes

Here we are referring to anyone in management at Hughes who was not directly on the manufacturing or testing floor. This includes the reporting hierarchy from Division Manager down to Frank Saia, the assistant manager for hybrid production. Hughes Aircraft (and thus all management) had a large stake in keeping the production and shipping of hybrid microchips at peak performance.

A primary problem for them was the 10% failure rate when chips were tested. This was a dramatic loss of money and time. Reducing this loss was important to Hughes, as was maintaining its pace of production. Changing production to produce chips that failed less often would reduce its pace. But "baby-sitting" the chips through the testing procedure to make sure they passed the tests (if at all possible) was a relatively easy way to reduce the loss from failed chips, particularly chips that were needed immediately by customers. Thus was started the process to declaring some chips "hot parts" that were to be rushed through testing as quickly (and with as low a failure rate) as possible.

One other important level of upper management was the personnel office. On more than one occasion, Goodearl spoke with people in the personnel office about her concerns and about the harassment she thought she was receiving because of her complaints about test skipping. But personnel did not have any independent way to investigate these incidents. Instead, on one occasion Goodearl saw the personnel officer walk directly from meeting her into office of Frank Saia, the manager of her division. Saia was one of the people she had complained to personnel about, and he was almost immediately notified of the complaint and almost as quickly summoned her to his office to shout at her. Thus personnel saw itself as on the side of upper management against workers who complained.

•Quality control

Ruth Ibarra (who's later married name was Aldred) was the primary Quality control employee with responsibility to oversee the testing environment for strict adherence to the testing rules. But there is nothing in the extensive legal documentation in the case that suggests she had an independent way to enforce her oversight. Time and again, we find that when she sees a problem, she has to report it to Goodearl or LaRue. Thus, Hughes did not really have an independent Quality control office, but instead required these individuals to report to people in charge of the testing of the product.

•Supervisors of testing

Both Goodearl and LaRue were floor supervisors on the testing area. Goodearl reported to LaRue and was designated to replace him when he retired. LaRue, however, was the one with the close relationship to the next level up of management, Frank Saia. And it was LaRue's agreement with Saia (based on pressure from upper management) to "baby-sit" parts that opened the door to begin skipping tests systematically.

The move to "baby-sitting" chips through the tests made it easy to take the additional step of skipping tests in the name of speed and efficiency. The pressure from upper management was relentless to ship the "hot parts," and middle management (in the persons of Frank Saia and Donald LaRue) were left to sort things out as best they could. This became a prime breeding ground for fraud done by middle management as a way to achieve objectives set by upper management.

Goodearl, however, was relatively new on the job and did not have the commitment to management values that LaRue did. This made it (somewhat) easier for her to question the skipping of tests and to take action by making internal complaints. When it became clear that her internal complaints were being ignored and that her job was threatened, she took the risk of blowing the whistle by reporting the incidents to the Defense Department Inspector General's fraud hotline.

•United States Government

The various agencies of the US government involved in this case include congress (via defense appropriations and law passed regarding whistleblowing protection), the Defense department (both as a customer to Hughes and as a watchdog over Hughes in the form of the Inspector General) and the Department of Justice (who joined the qui tam civil suit late in the game).

•The public

Certainly one interested party in this case is the United States citizenry. Their taxes were being used to fund the programs that Hughes was contracted to fulfill. In addition, if some of the military equipment using the chips malfunctioned, citizens might also be endangered. But the public are more a passive recipient of (potentially fatal) influence in the socio-technical system and not active participants.

•Members of the armed forces

Like the public, members of the armed forces are more passive recipients of influence from this socio-technical system. But they are more likely to suffer fatal consequences.

•Procedures

It is clear that the procedures for testing the chips and for dealing with complaints were the major thing at fault in this case. It was easy for LaRue and Saia to manipulate the system to ship untested chips, Goodearl and Ibarra had little protection when they made their complaints, their complaints were badly mishandled by management, and the Inspector General's office took an extraordinarily long time in moving to trial and getting a settlement.

One of the responsibilities of management is to design procedures that assure that the obligations of the organization are carried out appropriately. In addition to designing procedures that work management also sets the moral tone in the organization for how those procedures should be implemented in practice. Hughes failed in both these responsibilities.

•Documentation Procedures

The documentation in this case consists mostly of the "travelers" that went along with every chip. These were forms that specified how each chip was tosw be tested and that allowed documentation to accumulate (by initials, checks, etc.) that the tests had in fact been carried out and that the chip had passed. These paper documents were easily altered by management to make it look as though chips has passed tests that they had in fact failed. The only undeniable evidence in the Hughes criminal trial was a photocopy of a traveler saying "failed" on it that could be compared to the altered document in the files that showed "passed" written over erasure marks. Goodearl and Ibarra had made the photocopy knowing that LaRue was going to alter the traveler. This provided them with clear proof.

It is tempting to think that some sort of electronic monitoring system (using bar codes or some other information bearing medium) could be implemented that would make this sort of fraud by erasure impossible. But someone, likely someone in management, will still have passwords and access to the databases in this system and can still alter data. It is instructive to note that most financial loss from computer systems in banks and other organizations is from fraud perpetrated by those with trusted access to information. No matter how secure the technical part of a system is, there is no security if the personnel cannot be trusted.

So, if management wanted to alter information to make the chips appear to pas inspection, they would do so regardless of the kind of technical system was in place. What is needed, in fact, is independent supervision of the process.

•Oversight Procedures

The procedures in Hughes for handling oversight were badly flawed. Not only was there no independent oversight of the testing process, there was also no independent way to investigate an internal charge of fraud. Hughes did have a quality control group to check on the testing procedures, but their only course of action if they discovered something wrong was to ask the very people who might be doing the thing they were complaining about. Thus, Ibarra in quality control could only ask LaRue why such-and-such a test had been skipped, and often LaRue simply replied "none of your business."

LaRue's immediate subordinates were the "girls" on the testing line who were actually performing the tests. On several occasions, these people complained to Goodearl that LaRue was asking them to mark chips as passed when they had actually failed. LaRue would simply explain that the chips in questions were special and required different treatment. Or sometime he would not explain at all and simply say "Do what I say." It was clear that the atmosphere of the testing floor was one of great power disparity, and the girls were simply to do as they were told and not question management orders. In addition to this atmosphere of absolute obedience, was the lack of training that the girls had. They did not have, and were not given the chance to acquire, the information that LaRue had about the chips. This meant LaRue could simply ignore their questions and concerns.

If you look at the guidelines for ethical dissent you will find that one of the main sources of power in an attempt to dissent from an organization's decision is the information to prove one's case. This power was systematically kept from the "girls" of the testing floor.

In addition, Hughes did not have an independent way of investigating allegations of fraud or harassment when they were brought internally. The first reaction of those in the personnel department was to alert Goodearl's management that there was a problem. This might seem reasonable at first blush, but if Goodearl is, in fact being harassed, this makes it quite easy for management to punish her for reporting it. And in fact, this is what happened.

Several things can be said in defense of Hughes procedures in the socio-technical system. It is true that the "girls" did not have the knowledge to challenge LaRue's insistence that they pass a chip. It also true that they did not need detailed information to do their jobs, and that hiring people with minimal qualifications helped Hughes keep its costs down. In addition, it was not a widespread practice at the time to have independent oversight, either for quality control or for harassment charges. It was in fact cases like those at Hughes that have driven the concern for independent oversight and investigation.

•Laws and Regulations

The specific laws mentioned in this case are the False Claims Act, Whistleblower Protection Act, and the 1987 Department of Defense Authorization Act. These are discussed in brief in our section on US Whistleblower Law, These laws are under constant revision in the United States, and vary widely from country to country. If you are considering blowing the whistle, we recommend you find local legal expertise familiar with whistleblowing law. Our links page can help you look for this.

In addition to laws about whistleblowing, the reason Hughes got into trouble was by violating the Military testing standards required for hybrids. These were specified in great detail, and varies from chip to chip that was being manufactured. Hughes took advantage of the variation in testing standards from chip to chip by claiming (falsely) that some chips should skip particular tests. The complexity of standards was required by the variety of chip architectures, but it made enforcement of the testing standards more difficult.

•Data or Data Structures

The lot travelers that were attached to each chip are the most important data structures in this case. The data recorded on them were detailed lists of every procedure that was required, and a place to mark when it was completed. These lot travelers structured the testing regimen for each chip and documented that it had been followed. But they were easily subject to fraud, as evidence by LaRue's documented alteration of two AAMRAM lot travelers. Since the fraud is really a part of the procedure in testing, however, we will discuss the lot travelers in procedures.

Ethical Reflections

This document presents an overview of the ethical issues associated with the whistleblowing case of Hughes Aircraft. If we use the framework from chapter 2 the Hughes case highlights some important issues in five of the seven columns of ethical issues defined by the framework. In addition, some of these issues need to be addressed as more than simple individual ethical decisions about whether to blow the whistle or not; we will need to look to the group and national level issues involved.

The framework approach to ethical analysis was devised by a panel of ethicists, computer scientists, and social scientists. The point is that any particular computing system can be analyzed from both the perspective of social analysis and of particular ethical issues. The grid you see below was designed by the panel to serve as an analytic tool in thinking about any system. The idea is that each of the ethical issues can be analyzed at each of the levels of social analysis. If you click on the colored cells in the framework, you will be linked to a discussion of that ethical issue in the Hughes case.

This case is primarily about the extreme case in ethical dissent, whistleblowing. There is no column among the ethical issues for whistleblowing, but the ImpactCS approach allows us to see the complex issues of whistleblowing based on its component parts. It frames the ethical issues associated with whistleblowing as a complex mix of different ethical issues at several different levels of social analysis. Specifically, whistleblowing is about the use of power in service of some (ethical or unethical) end. Goodearl and Ibarra cite their concern for the safety of military personnel, but there is also a basic issue about fraud which falls under the Honesty and Deception column. Thus, for Goodearl and Ibarra, whistleblowing was about getting a powerful ally on their side in their struggle to influence Hughes Aircraft to properly test their chips.

Use of Power

Whistleblowing is usually analyzed in terms of balancing the duties the whistleblower has to the employer and to the public. But this approach isolates the whistleblower as the sole responsible actor. In actuality, whistleblowing takes places in a system. The employer, the employee, and outside agencies are played off against each other. Those who attempt to do the directing include, of course, the whistleblower, but also the employer, who will use both their own power and recruit that of others to strengthen their own position.

To help frame the use of power in this case, it will be useful to distinguish various sources of power. Raven (1993) lists seven different types of power, each with some overlap into others. This classification helps us to recognize both the sources of power that actors in the case have and the fact that power can be exercised even by the party that seems objectively "weak" in a situation.

The seven different types of power that Raven (1993) outlines are:

1. *Reward*. The ability to give or to withhold rewards from someone. Goodearl's placement in the testing operation was a reward to her for her earlier good performance.

2. *Coercive*. The ability to compel action under threat of punishment. Certainly Hughes used its ability to punish to attempt to influence Goodearl and Ibarra. Goodearl and Ibarra blew the whistle to recruit an outside agent (the Inspector General of the Department of Defense) who had the power to compel action by Hughes.
3. *Referent*. Power based on people desire to emulate an admired person. Classic examples are religious leaders, but in Hughes, both Donald LaRue and Frank Saia had some referent power because of their long service at Hughes. They were respected employees.
4. *Legitimate*. Socially sanctioned power, usually held because the person occupies a role that has responsibilities and associated power. Because of its position as the employer, Hughes had power to organize its affairs and to structure the jobs of its employees. The Inspector General had legitmate power because of its establishment by congress.
5. *Expert*. Power based on expertise. Because of both his experience and his education, Frank Saia had expert power. "The girls" doing the testing on the shop floor knew only their own station's routine, and so had little expert power.
6. *Informational*. Power based on information to which one has access. One can have informational power without being recognized as an expert in an area. For instance, some of "the girls" had information about how fraud was committed in the testing process (because they witnessed it, and had enough knowledge about the process to recognize it). But in its criminal trial, Hughes attempted to defend LaRue's actions based on his expert knowledge about what chips needed what sort of testing.
7. *Connectional*. Power based on who one knows or who one can contact. Goodearl and Ibarra were exercising connectional power when they called on the Inspector General to investigate Hughes.

People usually think of power only in the first two senses (reward and coerce). But all the different types of power can be seen operating in the Hughes case, and for each use of power we can ask the question "is this use of power ethical?"

To ask this question appropriately, we need criteria and procedure. The approach we present in chapters 4 and 5 is to use relatively straightforward tests such as:

- *Harm/Beneficence*: Does it do less harm or more good than the alternatives?
- *Publicity*: Would I want this choice published in the newspaper?
- *Reversibility*: Would I think was a good choice if I were among those affected by it?
- *Code of Ethics*: How does this choice stand in relation to the professional ethical standards of my profession?
- *Feasibility*: Can this solution be implemented given time, technical, economic, legal, and social considerations?

Each test can help understand a different facet of the ethical issues in a case, and the systematic use of the tests requires some knowledge of the structure and philosophy behind each test.

To do each of these tests on each of the uses of power for each actor in this case would be prohibitive here. But we will selectively use some of the tests from the perspective of each major actor in the case, and on several levels of social analysis (individual, group, national).

•Individual level

•Donald LaRue & Frank Saia.

LaRue and Saia were the first two links in the organizational hierarchy above "the girls" on the testing floor. They thus had legitimate power to structure the work of the environmental testing unit to best serve the purposes of the parent organization, Hughes. However, there were limits on their legitimate power that were imposed by employment law (e.g. non-discrimination) and by the government contract (e.g. what tests had to be performed, how the documentation had to be structured). This makes it clear that LaRue & Saia's legitimate power was really shared with other who also had legitimate claims. Did LaRue and Saia use their legitimate power ethically? In many ways, yes. They were given the power by Hughes in order to serve Hughes' end of making quality products while making a profit. Saia, for instance, changed the organization of the testing line to do the "gross leak" test early in the process, so that time would not be wasted on testing leakers that were easily detectable. LaRue carefully supervised "the girls" to make sure production quality and speed was maintained.

So where did they go wrong? The contract specified many things that needed to be done, and sometimes in what order. But it did not specify, for instance, that the "gross leak" needed to be done at a specific time, and Saia was free to move it around. But when they began omitting tests or falsifying tests, these actions clearly crossed the line. They did not simply go beyond the legitimate power. They actually disregarded the legitimate power (in their obligations in the contract) that others had over them. This clearly fails the reversibility test, since we want others to keep their promises to us, and so we should keep the promises we make to others. It also alerts us to the fact that we are dealing with obligations to respect the relationships and roles we have arranged with each other. So, they were overstepping the bounds of their own legitimate power and doing things that did not respect the legitimate demands of others. They were doing things (changing procedures) that looked a lot like things they had legitimate power to do, but they were imbedded in a web of obligations, and got the balance among them wrong.

In addition to violating their obligations in the contract, they also may well have been putting lives in jeopardy because they were allowing inadequately tested chips to be shipped out to be used in military hardware. This clearly fails the harm/beneficence test. They would have been wrong to skip the tests under most any circumstances, but especially wrong when the skipping could result in unsafe systems and potential loss of life. Goodearl and Ibarra cite their concern about safety as the primary reason for their

whistleblowing. There were less upset because LaRue and Saia were merely bending the rules, but were driven to action by the safety implications of those departures from the rules.

LaRue and Saia also used the coercive power that came with their position to punish Goodearl because she was not being what they called a "team player." Surely this fails the reversibility test, and also the publicity and harm/beneficence tests. It is important to recognize why it fails these tests. Every organization wants employees to be team players. This expectation is reasonable, and it may be necessary to sanction or even dismiss those who do not meet it. But the code of ethics test helps us to see how the balance was missed. The ACM Code of Ethics. Section three of that code specifies organizational leadership imperatives: the responsibilities of those who are leaders in organizations that deal with computing. Item 3.1 reads:

ACM member and an organizational leader, I will articulate social responsibilities of members of an organizational unit and encourage full acceptance of those responsibilities.

So, one of the responsibilities of an organizational leader, according to the code, is to make clear to employees their social responsibilities and to support employees in accepting those responsibilities. But LaRue and Saia were only emphasizing loyalty to the organization, to the team. And not only were they not "encourag[ing] full acceptance of [social] responsibilities" they were actively punishing Goodearl and Ibarra for attempting to fulfill their social responsibilities.

Finally, LaRue and Saia used their expert power to claim a right to selectively test chips. On the surface, this seems reasonable and perhaps even laudable. If, in fact, their use of their position as experts on testing to make sure the chips were make the line more efficient or even to simply make more money for Hughes while maintaining the required quality, all would be well. But in fact, they misused their expert power either to claim (falsely) that no harm was being done or to minimize the harm being done by the fraud they were committing. In either case, this clearly fails the publicity test. The publicity test focuses on personal character (what would publicity reveal about the kind of person you were?) and were are thereby alerted to the possibility that their misuse of expert power is wrong because it highlights the lack or a virtue (honesty) or the presence of a vice (greed, perhaps, or cowardice). Since honesty and deception is one of the main ImpactCS categories, we will deal in more detail with this angle of LaRue and Saia's unethical behavior in that section.

•Goodearl and Ibarra

In her initial attempts to reform the testing procedure from the inside, Goodearl attempted to use her informational/expert power to convince the organization that it needed to follow the testing protocol. She was stymied in these attempts and punished for her efforts. If we look at the IEEE guidelines for ethical dissent, we can see that Goodearl blundered ahead through the steps of dissent without taking time to make her case. Thus,

though morally right, her tactics were flawed. The IEEE guidelines suggest that a careful case be put together, based on the best information, that helps to maximize the goals of both the dissenter and of the organization. Goodearl spent most of her time simply telling other that they had to follow the rules. She was correct, but tactless.

Here the IEEE guidelines and the ethical tests we propose converge. The feasibility test (the last item in our list of ethics tests) ask questions about how best to proceed in attaining the ethically desired goal. Can it be attained at all? If a thing cannot be done, one is usually not held responsible for not achieving it. In the criminal trial, Hughes was found guilty of fraud, but LaRue was exonerated because the jury felt he was put in a position in which he had few options except to do the organization's bidding. On the other hand, one of the jobs of a thoughtful ethical dissenter is to form coalitions with others in the organization who can help to support the dissenting case.

Goodearl in fact attempted to do this by joining forces with Ibarra in quality control. Ibarra has expert power in the organization, and could help to strengthen her case. But although it appeard that Ibarra had power, the organizational reporting schemes required her to route all her complaints through Don LaRue, the very person who was committing the fraud. So, it appears that the organizational setup was one that gave the appearance of independent review, but actually gave no power to the independent reviewers.

For this reason, Goodearl and Ibarra both ended up using connectional power by establishing outside connections with the General Inspector's office. This outside agency had power that was independent of the immediate reporting chain in which Goodearl and Ibarra were stuck. It may have been possible for Goodearl and Ibarra to still stay inside the organization and find some ally outside their immediate reporting chain. In fact, Goodearl attempted to do so by speaking with the Personnel office. But again, the reports were quickly funneled back into the immediate reporting chain, this time to Frank Saia, LaRue's direct supervisor.

Still, it seems apparent that Goodearl (at least) and Ibarra (possibly) were less than strategic in their attempts to find an inside remedy the problems they saw. This is certainly excusable in that they were not professional computer scientists or engineers, and did not have the broad knowledge (or expert power) required to critique the system and play it against itself.

If we approach their decisions from the viewpoint of the ethics tests, we can see that they felt they were helping to reduce harm to military personnel and that this overweighed the harm they might do to themselves and their families or to the organization. They still claim that they would do the whistleblowing over again. But they did pay high personal price.

In terms of the reversibility test, were they treating their supervisors and Hughes with appropriate respect? Surely they were doing things that hurt Hughes and their supervisors, but they could be done in a way that respected their rights. In the criminal trial, it became clear that there were personal animosities between Goodearl and LaRue,

and it seemed that they were asking "the girls" on the floor to pick sides in the ensuing fight. So, there is some evidence that Goodearl was not respectful of LaRue. But again, this may be excusable given the severe pressure she was under.

Certainly Goodearl and Ibarra are celebrated as virtuous persons because of their whistleblowing, and this make the publicity test easy to pass. An important point to remember here is that Goodearl does not need to be sainted, to be perfect in virtue, in order for us to say that she showed courage. In fact, in the criminal trial, it became clear that Goodearl had a variety of character flaws (e.g. bragging about untrue exploits, making immodest claims about personal achievement) and the defense for Hughes used these flaws to their advantage in attacking her credibility. But when we step back from the case it is easy to say that both she and Ibarra showed extraordinary perseverance and courage in pursuing their whistleblowing claim.

•Group Level

In thinking about the ethics of LaRue and Saia's actions we have already talked about Hughes legitimate power to set up a system that achieved their goals. There we mentioned that authoritarian atmosphere that Saia and LaRue maintained in the environmental testing unit. But the atmosphere was spread more widely throughout Hughes. When Goodearl attempted to report incidents to the Personnel office, her complaints were not taken seriously, but immediate funneled back to her superiors.

It seems, then, that although Hughes had legitimate power to structure its work environment so that it could achieve its corporate goals, it failed in its duty to "articulate social responsibilities of members of an organizational unit and encourage full acceptance of those responsibilities" (ACM Ethics Code 3.1). This broader failing worked, in the end, against even Hughes' own goals, since they were found guilty of fraud in a federal criminal court.

We can begin to construct a case here for the ethical responsibilities of organizations to encourage ethical behavior in their employees. We can base this case on the various ethical tests we have been using, as long as were are willing to say that an organization can show virtues, weigh harms and benefits, and show respect.

•National Level

Congress certainly used legitimate power in its design of laws to encourage whistleblowing. The original law that supported Goodearl and Ibarra was passed during the civil war to encourage (and reward) those people who would sue, on the behalf of the government, perpetrators of fraud. These laws are under regular revision, and Congress sees the laws as supporting an environment in which employees are encouraged to think about the ethical issues associated with their behavior as agents of organizations.

The Inspector General's office is authorized to use coercive power (through lawsuits and criminal proceedings) to compel cooperation on the part of government contractors. In the Hughes case, most of use think this is a reasonable use of power. But if you go to the

Inspector General's reporting hotline web page you will discover that the office has wide discretion about how much they will investigate anny particular allegation. Feasibility is surely one of the issues here – the office can only investigate so many incidents at a time. But in additon there is an issue of when they have enough evidence to justify an investigation or a suit. In Goodearl and Ibarra's case, the two employees were encouraged to collect evidence that the fraud they alleged was actually being perpetrated. The point here is that even the investigative powers of a legitimate authority have limits. Describing those limits takes us outside the scope of this analysis, but is a useful analysis. In attempting this analysis, remember not to reduce the moral and ethical responsibilities of the Inspector General to the legal requirements.

Safety

In the Therac case in this chapter we describe the safety issue associated with the implementation of a computing system in a real sociotechnical system. In the Hughes case, the safety concerns occur at the manufacturing level rather that the implementation level. This brings home the point that a sociotechnical system needs to include those systems that contribute to ite manufacture. And it highlights the ethical responsibilities of computing professionals to design systems that take into account the way components are designed in the real world.

Of course, it is impossible to design a system to avoid fraud on the part of component suppliers. But one can certainly think about the needed redundancy to make a system work even if some of its parts fail. The estimates of this needed redundancy need to take into account the likelihood that all the parts that are delivered may not be up to specification. This can dramatically increase the likelihood of component failure and make redundancy more important.

Privacy

A clear privacy issue in this case is the handling of the reporting of fraud in the Personnel office. Personal information that was given in confidence was transmitted to the complainer's immediate supervisor, resulting in retaliation for the complaint. This is a dramatic compromise of privacy. The ethical issues here were handled in more detail in the section on power.

Equity & Access

Why were the testers in the environmental testing area called "the girls"? This was a term of endearment and the testimony at the criminal trial bears this out. But in addition, it was also a term that emphasized their lack of power. Was Goodearl "one of the girls"? Apparently not. But she was definitely subordinate to LaRue, who often treated her complaints with the same arrogance as those of "the girls." Goodearls' last supervisor was also female, but this female supervisor had learned the lesson that she must go along with the corporate culture in order to get along. Thus like in many organizations, the women who advance most quickly are those who reflect the male corporate values.

When is it dishonest to skirt the rules? Before answering this "never" remember that unions have perfected the "work-to-rule" slowdown as a form of collective action against management. These work slowdowns are accomplished by simply adhering literally to every rule for repair, clean-up, paper filing, etc. that has been established. It does significantly gum up the works. So in fact, one has to skirt the rules to get an organization to function at all.

There are clear cases of fraud, like the one perpetrated by Hughes, and we can condemn these straightforwardly. But between outright fraud and a work-to-rule slowdown there is a gray area in which people of goodwill can disagree. Labeling some of the chips "hot parts" and expediting their testing, or even babysitting them through the process, seems to be a legitimate bending of the rules. For each of these, using the ethics tests we have presented here help in establishing the important issues. The tests do not always make the answer clear, but they help to make the questions more clear.

Part of Goodearl and Ibarra's dilemma was that if they became "team players" according to LaRue and Saia's rules, they would be participating in the fraud the Hughes was committing. The doctrine of respondent superior helps some in this case, in that it holds the employer responsible to the action of the agent when the agent is really acting on the behalf of the employer. This is, in part, the reasoning that the jury used in not convicting laRue of fraud, but convicting Hughes. Still, this legal rule does not let the agent entirely off the hook in the legal system, not does it address the moral responsibility of the agent. Goodearl and Ibarra knew that people would be placed at risk if the chips were not properly tested, and they decided that this risk outweighed the interests of Hughes.