Data Analysis “Back in the Day”: The Early Career Experiences of Nine I-O Psychologists

The availability of the personal computer (PC), statistical software, and the Internet has had undeniable effects on I-O psychology. Without such technological advances, for instance, there’d be no virtual teams, no computer-adapted testing, and no cyberloafing. To better appreciate the impact of technology on the current state of our discipline, it’s helpful to reflect on the technology used in the recent past. In preparing this installment of the History Corner, we interviewed nine seasoned I-O psychologists: Terry Beehr, Ilene Gast, Lawrence Hanser, Milton Hakel, Norman Peterson, Susan Reilly, Neal Schmitt, Paul Thayer, and Lauress Wise. We asked them each to describe the technology they used during their early careers to conduct data analysis, and we asked them to reflect on how technological changes have affected the way in which I-O psychologists conduct research. In the following sections we discuss how calculators, early computers, and PCs were used “back in the day” to conduct data analysis. We then discuss how I-O psychologists wrote their research reports prior to the advent of PCs and word processing programs.

Conducting Statistical Analyses Using Calculators

Many of the interviewees told us that they conducted statistical analyses using calculators, especially for small datasets and for class assignments. Thayer told us that he conducted his doctoral work in the early 1950s in Dr. Herbert Toops’ lab at Ohio State University. Toops had a mechanical hand-crank calculator in the lab that looked like a typewriter and had a crank that the user would move forward for addition and multiplication and backward for subtraction and division. The psychology department had a calculator lab, with about 20 machines that graduate students could use for their research. Statistics classes would often have lab sessions in the calculator lab, and Thayer remembers his fellow students having races to see who could do their calculations the fastest. According to Thayer, the calculators at Ohio State University were of the Marchant brand (another common brand was Friden). A pic-
ture of a Marchant hand-crank calculator is shown in Figure 1. Using the calculator was laborious as there were often mistakes in the calculations and data entry, which required the user to start over. These calculators had no memory and no printout; thus, there was no record of what took place other than what the user wrote down.

**Figure 1.** This is a Marchant H9 Calculating Machine. The user would input the numbers for the calculation using the keyboard; the numbers that were entered would appear in the row of nine dials in the upper right corner of the machine. There is a crank on the right side of the machine that was used to conduct the calculations. The user would rotate the crank forward for addition and multiplication and backward for subtraction and division. The results of the calculation would appear in the row of 18 dials directly above the keyboard. (Image is from http://americanhistory.si.edu/collections/search/object/nmah_690715 and appears courtesy of Kenneth E. Behring Center, Division of Medicine & Science, National Museum of American History, Smithsonian Institution.)

An early type of calculator—the “four-function calculator”—could only perform four mathematical functions: addition, subtraction, multiplication, and division. Many statistical and psychometric equations require calculating the square root of a number. Hanser told us of an iterative algorithm he knew of to get an approximation of a square root, which is illustrated in Table 1. There were other tricks of the trade for simplifying statistical computations. Wise, for example, told us that it was possible to get the squares and sums of squares in a single step using a Marchant calculator by entering the data as a nine-digit string (e.g., to obtain the sums of squares and crossproducts for the numbers 4 and 10, enter 004000010 and square this number to obtain 16,000,080,000,100, or 16[4²], 80[2×4×10], and 100 [10²]).

**Table 1**

<table>
<thead>
<tr>
<th>Step</th>
<th>Procedure</th>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Start with a guess of the square root of .92^2</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>1a</td>
<td>Divide .92 by the guess</td>
<td>.91</td>
<td>.979</td>
</tr>
<tr>
<td>1b</td>
<td>Obtain the average of the divisor and the result</td>
<td>.979 + .94</td>
<td>.96</td>
</tr>
<tr>
<td>2a</td>
<td>Divide .92 by the average</td>
<td>.960</td>
<td>.958</td>
</tr>
<tr>
<td>2b</td>
<td>Obtain the average of the divisor and the result</td>
<td>.960 + .958</td>
<td>.959</td>
</tr>
<tr>
<td>3a</td>
<td>Divide .92 by the average</td>
<td>.959</td>
<td>.959</td>
</tr>
<tr>
<td>3b</td>
<td>Obtain the average of the divisor and the result</td>
<td>.959 + .959</td>
<td>.959</td>
</tr>
</tbody>
</table>

The introduction of square-root hand-held calculators represented a major convenience for researchers. In fact, Beehr received one as a Christmas gift in the 1970s; it cost approximately $100. Gast told us about her experiences using HP statistical calculators in the 1970s, such as the one shown in Figure 2. The group she worked in only had one of
these statistical calculators. It was able to run regression and other statistical analyses and used magnetic strips, called cards, which held programs or data. This was a big advancement from her days as an undergraduate at American University, which only had a single four-function calculator and a waiting list to use it.

**Calculating With Computers**

When analyzing larger datasets, using a calculator was often impractical; therefore, researchers often used mainframe computers. Most often the original data for a study were collected on paper and had to be loaded into the mainframe computer’s memory for analysis. In addition, the syntax for running the statistical programs on the mainframe also had to be inputted. It was not possible to directly input the data and syntax into a computer. Instead, punch cards, also known as Hollerith\(^1\) cards, were often the primary medium for inputting data (see Figure 3).\(^2\) These were small—often 7\(\frac{3}{4}\)” by 3\(\frac{1}{4}\)”—sheets printed on cardstock that contains 80 columns and 12 rows. Data were saved on the punch cards by punching out small squares in

---

\(^1\) Hollerith, T. J. (1859–1918) was an American inventor known for his punch cards, which were used for data input and processing in early computers.

\(^2\) Punch cards are data storage devices consisting of perforations, which can be read by a machine to represent data values.

**Figure 2.** The top image is of an HP-65 calculator. The middle image shows a black card that is partially inserted in the calculator. The card is fully inserted in the bottom image, it appears above the row of keys labeled A-E. This card could store programs and data. HP sold packages of cards for different purposes, including the two statistical packages (e.g., normal distribution, correlation, analysis of covariance). (Images appear courtesy of Nigel Tout, Vintage Calculators Web Museum, www.vintagecalculators.com.)
each column, much like the infamous Flori-
da voting machines containing the hanging
chads in the 2000 presidential election.
Thayer remembers hand punching data
for his dissertation using a small handheld
machine, such as the ones shown in Figure
4. He said it was a good practice to veri-
fy one’s work by placing the card back in
the machine and repunching it. Mistakes
could be identified by looking at the card
to see if more than one hole was punched
in each column (the importance of ensur-
ing the cards were error free is described
below).

Over time the, small handheld machines
were replaced by larger keypunch ma-
hines. These machines were about the
size of an upright piano (see Figure 5 for
an example) and were run by keypunch
operators, which quickly became an occu-
pation. Most universities contained rooms
Figure 4. This figure contains three images of manual punch card machines. The first bears a strong resemblance to that designed by Hollerith. The second machine is the Wright Line manual card punch and the third is the IBM Type 11 electric keypunch. (The first image is courtesy of Wikimedia commons and the second and third images are courtesy of Computer History Museum.)
of keypunch machines that graduate students could use. For those students with grant funding, keypunch operators could be hired to do the actual work. The keypunch machines contained a keyboard, hopper for new cards, and an output stack for punched cards. A keypunch operator would enter the data using the keyboard, much like data entry is conducted today on a computer. However, as the data were typed, holes would be punched on the punch card. After 80 characters were entered, the card moved to the left (where it could be inspected) and another card was fed in from the hopper. After punching the cards, it was often a good idea to make a second copy; this could be done using a duplicating feature on the more advanced keypunch machines. Some historical footage providing more information on keypunching are available on YouTube (see https://www.youtube.com/watch?v=oaVwzYN6BP4 and https://www.youtube.com/watch?v=YXE6HjN8heg).

After the cards were punched, they had to be inputted into the mainframe computer. This is the most time-consuming part of data analysis “back in the day.” Most campuses only had one mainframe computer and everyone on campus had to share it, not only researchers but also administrators using it for processing payroll and grades. A researcher would take the cards to the computer center and turn them over to the computer center staff, often by placing the deck of cards in a metal tray. Then the wait began. The cards were put in a long queue of various jobs for the mainframe. After hours—or sometimes days—of waiting, the cards would be fed into the machine and the mainframe would conduct the data analysis. The analysis usually went pretty quickly—it was the backlog of jobs that the single mainframe computer had to process that took time. In addition, each job’s cards had to be manually carried to and from the mainframe by a computer operator. The results of the analysis were outputted on paper and both the paper and cards were later collected by the researcher. Sometimes the entire process from dropping off the punch cards to obtaining the printout with the results could take 24 hours (in Gast’s experience) or even 2 weeks (in Beehr’s experience).

If everything went as planned, the analysis was complete and the researcher could begin interpreting the results. However, things did not always go as planned. If
there was a mistake in the syntax code, then the printed output might reflect this (much like what occurs today using modern statistical software). At other times, the researcher would only receive a printout stating “JCL [job-control language] error.” At this point, the researcher would have to determine what the mistake was, repunch a portion of the cards, and then head back to the computer center and wait. As a result, an error could cost the researcher hours or even days. This is the reason researchers would spend much time double-checking their punch cards and thinking carefully about their analyses and syntax. As several interviewees pointed out, you could not simply play around with different analyses like some researchers do today. It simply was too inefficient and time consuming.5

Punch cards were problematic for other reasons. If a card was torn or bent it would have to be repunched. In addition, cards sometimes became jammed in the mainframe computer or other card processing machines and had to be replaced. Hanser remembers having to use a card saw (a special thin knife without a handle) to saw through a set of jammed cards in a card sorting machine. Researchers walking across campus were always fearful of their cards being dropped or blown away by the wind. Hakel recalls some of his colleagues numbering punchcards and having to check the order of the cards after they were returned by the computer operator.

Eventually, the process of using punch cards to provide commands to the mainframe was replaced by computer dummy terminals. One such terminal was the DECwriter, a combination keyboard and dot matrix printer (see Figure 6). The dummy terminal was not a computer itself but was instead used to remotely control a mainframe computer (such as the IBM 360 or 370) via an acoustic coupler (an early dial-up modem). The syntax code could then be entered using the dummy terminal, and after the job was completed, the output would be printed on the dummy terminal. The process is akin to a virtual punch card and dummy terminal “Mad

![Figure 6. The IBM DECwriter, which looks like a cross between a dot matrix printer, keyboard, and typewriter, was used as a dummy terminal to control a mainframe computer. (Image courtesy of Wikimedia commons.)](image)
Men style” GoogleSearch that you can try for yourself online (http://www.masswerk.at/google60/).

Larger datasets could also be stored using magnetic tape (often originally created by reading a stack of punch cards). Thus, a researcher would specify which reel of tape was needed. The tapes were usually not handled by the researcher; instead, they resided in the mainframe computer center’s tape library and would be loaded onto the mainframe by a computer operator. Although magnetic tapes were more stable than punch cards, they were not without problems, as is evident by two stories Hanser told us. Once he was running data using magnetic tapes in a trailer at a military post and a wire bouncing against the outside of the trailer caused the data on the tapes to become scrambled. Sometimes the tape itself would be physically damaged. In these situations, one of his colleagues, Frances Grafton, painted a compound called “magnaflux” onto the tape to visually reveal where the magnetic bits of data were (which could be seen because the data were not packed very close together on the tape).

Around this time, many large organizations had their own mainframe computers. However, those that did not had to lease time on a mainframe computer. At universities, computer time could be charged to a grant or to the department. I-O psychologists at organizations that leased mainframe access had to worry about the cost of making mistakes with their analyses and the length of time it took to run more intensive analyses and larger datasets, each of which could cost hundreds or thousands of dollars.

Statistical software. Psychologists conducting data analysis prior to the advent of point-and-click statistical software (e.g., in the late 1980s to early 1990s) had to be versed in syntax programming for several different software packages. In the early 1950s, a researcher had to review the formulas for a particular analysis and then think about how best to program them into the mainframe. Later, statistical packages such as COBOL (short for COmmon Business Oriented Language) and FORTRAN (short for FORmula TRANslation) became available (in 1959 and 1957, respectively). COBOL was good with processing data (e.g., merging, sorting), whereas FORTRAN was better for statistical analyses and computation. Later, more powerful statistical software was released. Several of the psychologists we spoke with used BMDP (short for Bio-Medical Data Package), which was originally developed for the biomedical field in 1965. Beehr used OSIRIS (short for Organized Sets of Integrated Routines in Statistics; Van Eck, 1980) at the University of Michigan. Another used P-STAT, a program originally developed at Princeton University that earned the distinction of being called the “statistical package that doesn’t mess around” in PC Magazine (Ramsay, 1989, p. 130). SAS and SPSS became available in 1966 and 1968, respectively, and eventually became the most prevalent statistical packages used by I-O psychologists. Ofentimes, psychologists used whichever
system (e.g., BMDP, SAS, or SPSS) to which their employer or university had access, which meant learning a new statistical package when you changed organizations.

Writing Reports

After the data analysis was complete, I-O psychologists often had to write up the results. Today most of us do our writing sitting at a computer; however, “back in the day” desktop computers and laptops with word processors did not yet exist. Just about everyone we spoke with handwrote the text of their theses and dissertations. Most then paid a typist to type the text onto paper using a typewriter because most researchers of the time were not skilled in touch-typing. Using a typewriter made it very difficult to revise text. Sometimes the text could be changed with white out or by actually cutting and pasting the paper itself. More substantial changes might require an entire section to be retyped. Most professors were aware of this, as well as the financial state of their students, and often restrained themselves from asking their students to rewrite major portions of their text.

As technology advanced, typewriters were replaced with word processing machines, such as the Lexitrons (see Figure 7) that Reilly used for typing technical reports. Using this machine, it was possible to type, edit, and print a report. The Lexitrons also had a proprietary floppy disk for saving the report. It was also possible to have the text of a report placed onto punch cards and processed on a mainframe (as Gast did for a graduate school paper in 1974). Giddings and Zimmerli (1972) developed a program entitled Thesis 3.5 that some of Hanser’s classmates used for their theses and dissertations.

Advent of Desktop Computers

Needless to say, the arrival of desktop computers revolutionized data

Figure 7. This is a Lexitron word processor, model VT202 (image courtesy of the Computer History Museum).
analysis and report writing. However, when these computers became available in the 1980s, most organizations would have only a few computers per department. In other words, I-O psychologists did not have computers at their desks. Instead, they would have to wait until a shared computer became available. Data at this time could be stored on very large removable disks such as the Bernoulli disk, which contained 10 MB of data in a cartridge about the size of a ½ inch stack of letter sized paper. Smaller files could be stored on floppy disks (the most common sizes were 3½, 5¼, and 8 inches), and some computers used small magnetic strips for holding data. However, analyses on many of the larger datasets continued to be conducted using mainframe computers, especially if the dataset could not be stored on a floppy disk.

Desktop computers assisted greatly with writing reports, theses, and dissertations. As word processing software became available, it was no longer necessary to manually type text using a typewriter or a Lexitron. This made editing typewritten text much easier, as text could be copied and pasted without having to retype entire sections of a paper.

Modern Statistical Analysis

Eventually, desktop computer storage became adequate for storing large datasets and for running programs like SPSS and SAS. This meant that the days of walking across campus to the computer center with a pile of punch cards or using a DECwriter with an acoustic coupler to run a regression were over. Everyone we spoke to said that the reactions of their colleagues were overwhelmingly positive. It made data analysis much more efficient and flexible. In addition, collaboration with I-O psychologists who worked at different institutions became much easier. However, it also became easier for researchers to get by without fully understanding the math behind their statistical analyses or to sit down at a computer and run multiple tests and “fish” for significant results.

Slide Rules, Manual Factor Analysis, and Shortcut Statistics

Finally, we also asked the interviewees about slide rules, conducting factor analyses by hand, and usual shortcut statistics. Few, however, had experience with these. Many had used slide rules but not for their psychology work; slide rules were more common for high school and college courses, especially in trigonometry, chemistry, and engineering. Although most of the interviewees had heard stories of conducting factor analyses and rotations by hand, none of them were directly involved in this work (mainframe computers had made this task obsolete). According to Schmitt, Louis Leon Thurstone spent months doing factor analyses by hand and had papers with the analyses pasted over the entire walls of his office. Although some textbooks make note of obsolete statistical formulas that were used to save time (e.g., KR-21, the use of phi coefficients in lieu of Pearson correlations), the interviewees we spoke to said that by the time they entered the field, use of these statistical shortcuts was no longer necessary.
Summary

Technological changes have greatly impacted the way in which I-O researchers collect and analyze their data and write their research reports. By minimizing the “grunt work,” these changes have made the research process faster and more efficient. Perhaps a TIP History Corner article published 50 years from now will reflect on the technological limitations faced by researchers in the early 21st century.

Notes

1 Herman Hollerith invented an early punch card machine for use in the 1890 U.S. Census; his company was a predecessor of IBM (Aul, 1972).

2 Another option was to collect the data on optical answer sheets (e.g., scantron or bubble sheets) as was described in the last TIP History Corner (Cucina & Bowling, 2016).

3 This was especially the case during the day. Some of the interviewees told us that they would try to run their analyses during odd hours (e.g., over the weekend or in the middle of the night) as the turnaround time was quicker.

4 Peterson had experience working as a computer operator at an insurance company. He operated an IBM 1401 that read in punch cards and stored the data onto magnetic tapes. The tapes were then used as input (and output) for an IBM 7070 (discussed in our previous column, Cucina & Bowling, 2016) which was controlled using a teletype console (which often look like a DECwriter) and punch cards. Much of his work involved updating the insurance records on the tapes.

5 Peterson pointed out that if you had access to a lot of money, much of the “grunt work” could be contracted out. He said that there were students and other “guns for hire” who could keypunch your data, program your analyses, and handle the troubleshooting if you had money. However, most students lacked these funds and even some organizations hiring I-O psychologists balked at doing this.

6 According to Larry Hanser, Frank Medland of the Army Research Institute had developed a factor analysis program that could be run using a card sorting machine. It usually took all weekend to run the factor analysis.

References


SIOP 2016 has it ALL!