

A Brief Personal History of Media Optimization

The author of this document, Bill Harvey, was involved in many of the major developments in this field, and so this history is written from the author's personal perspective. However, a more complete picture would require interviews with others so as to round out the total history of media optimization.

The document was first written in the year 2000. It was updated in 2021, and we expect it will be updated again in the future.

It is a story of applying rigorous mathematical and logical approaches to a field that was more of an art than a science as our story begins. By 2021, the field of media planning and buying is a spectrum that ranges from art to science, with the most successful players tending to be on the science end. However, the practitioners on the art end have not necessarily observed that fact.

The idea of optimization derives from an engineering discipline called "Operations Research" and known as "OR". OR consists of a set of tools and approaches known respectively as "algorithms" and "heuristics". Algorithms are mathematical equations, while heuristics are fuzzier methods i.e. they are not equations. Both algorithms and heuristics are aimed at improvement in the operations of an organization.

"Optimization models" are a type of algorithm intended to provide the best possible solution to some problem facing an organization. Where the problem itself is so complex that finding the best possible solution could cost more than the benefit of doing so, the optimization models generally do not attempt to find the best possible solution, but instead seek to find extremely good solutions within reasonable cost and time parameters. This in fact is the more common situation. Although in the latter case what is sought is literally "improvement" rather than "optimization", these models are still conventionally called optimization models in all cases.

Media Optimization

The impetus to the application of optimization to media selection by advertisers and agencies came in the early 60s with the arrival of mainframe computers in the offices of the largest advertising agencies. These agencies found that they were initially utilizing the computer only for payroll, whereas the reasons for acquiring the computer included the belief that an agency could gain a competitive edge in

attracting new business, and keeping clients, by spending their advertising dollars more cost effectively using the computer. Thus began the development of media optimization models.

All media optimization models require the input of media audience and cost data for all measured media vehicles, and the input of the brand's requirements in terms of budget, target audience, reach/frequency, types of programs/ publications not acceptable for non-quantitative reasons, and other factors. The model is a complex set of equations which considers all of this input, and outputs one or more "best schedules" within the budget constraint. The parameter to be maximized (e.g. Target Audience exposures) is technically known as the "objective function". The optimizer is designed to maximize value by selecting those vehicles with the lowest cost in relation to whatever parameter is to be maximized.

The focus for the first five decades was almost exclusively on maximizing the target reach within a given budget, or determining the budget needed to attain a specific target reach, usually the former. Since television was the main medium, and television audience data were generally panel-based and initially not available at the respondent level, models to approximate reach were devised, sometimes using pair duplication data for all possible pair combinations among what was at the outset only a few hundred national television program series. As cable vastly increased the number of programs, the simpler "curve" method became the norm, in which reach is predicted from GRP/TRP.

The reach initially was typically four-week reach. Later when Erwin Ephron popularized the recency goal, some leading practitioners switched to weekly reach.^{1,2}

Within an incremental optimizer, one media vehicle is selected at a time, based on its having the lowest cost per incremental target viewer, after estimating the unduplicated audience that would be added to the vehicles already selected by each remaining vehicle option. Often the best choice is to add another insertion to a vehicle that has already been selected one or multiple times already. This reflects the fact that there will always be vehicles whose low CPM overcomes targeting considerations. In the real world, adding too many insertions into the same program series tends to limit reach and provide too much frequency. However, curve-based reach models are insensitive to the multiple-insertions problem and tend to over-estimate reach of multiple insertions tactics. To compensate for this

defect, optimizers have dashboards which allow restrictions on the number of spots placed in the same program series during the same episode/week/etc. (a heuristic).

The builders of media optimization models studied the way media planners and buyers conventionally selected media vehicles, and constructed their models to mirror these conventional procedures. In doing so, they sought to move from the heuristics being used by planners/buyers into the use of true algorithms instead.

Media Impact

Early modelers found that planners/buyers did not purely select vehicles with the lowest CPM Targets, nor could their selections be explained in terms of the audience and cost data supplied. When the modelers queried the planners/buyers to find out why the computer could not reproduce the media selections actually being made, it was discovered that planners/buyers were implicitly adjusting the quantitative data based on qualitative factors derived from their own experience.

For example, the early media models produced schedules using large amounts of radio and outdoor, two media types with low CPMs and low CPM Targets against a number of Target Audiences. However, in the real world, media planners had rarely used these media to such an extent, allocating much larger budget proportions to television, for example. This was because TV was believed to have far greater "impact" per exposure.

Unfortunately, there were no compelling data to support this belief. Instead, there were a number of one-time studies, often sponsored by the media themselves, showing conflicting results as regards relative media impact on commercial recall, sales, and other payoff measures. Studies sponsored by magazines, for example, tended to show magazine impact at parity with TV, while studies sponsored by TV tended to show TV having more impact than magazines, and so on.

In order for the modeling process to continue, the modelers requested that planners/buyers make their qualitative impact "hunches" explicitly quantitative in the form of "impact weights". A typical system would equate a primetime network TV 60 second commercial equal to 100, and then other media would be given lower weights in relation to that. For example, a primetime network TV 30 second commercial might be given a weight of 75, a four color "bleed" page in a magazine might be given a weight of 70, and so on.

The modelers argued that, at least, media planner/buyer judgments would be out in the open and would be employed in a consistent manner through quantification into impact weights. This convinced a number of agencies to create their own sets of impact weights.

Harvey's agency, Grey, had all senior and junior media personnel each create a personal set of impact weights across media types, and studied the results. The results showed that even among senior media personnel, there were great differences in the patterns of impact weights assigned to different media.

Grey concluded that any optimization results could be produced by manipulating the subjective impact weights, so that an optimization model could be forced to agree with the actual media schedules that had been selected without optimization. This caused Harvey to conclude that new media measurements would be needed to provide objective impact weights before the real value of optimization could be realized.

Pioneering Media Optimization Models of the Early 60s

The first two models were created by BBDO and Young & Rubicam respectively.

BBDO's Linear Programming or "LP" model solved a set of equations to produce a final schedule which provided the maximum number of Target exposures weighted by the judgmental impact scores that had been input. Non-linear elements such as frequency discounts on certain media, and audience duplication among different media and among successive insertions in the same media vehicle, were cleverly handled by means of built-in statistical adjustment factors that tended to limit the model's initial tendency to build up too many insertions in the media vehicles with the lowest impact-weighted CPM Targets—an approach which would tend to limit the unduplicated reach of a schedule.³

Y&R's High Assay Media Model ("HAMM") differed from BBDO's LP model by placing much more emphasis on the weekly performance of the media schedule. Magazines, for example, tend to accumulate their audiences over a period of weeks, and these patterns were estimated based on special tabulations of early Politz magazine audience surveys so as to deliver patterns of weight reflective of the seasonality of a brand's business.

The words "High Assay" in the name of the model placed emphasis on

the value of targeting, and reflected the notion that a small percentage of customers actually represent a large percentage of total business opportunity for any marketer—and that a marketer's top priority is to find ways to identify and discriminately reach these high value target customers. This idea is resonant today in digital and addressable TV, although controversies still exist as to the relative payback of targeting convertibles vs. loyals.

Harvey was able to study these first two models through accounts that Grey shared with BBDO (General Electric) and Y&R (Procter & Gamble). Seeking to improve upon these pioneer models for Grey, Harvey queried outside suppliers such as MIT and the British Information Technology company CEIR, among others, and studied their proposals.

The CEIR model was called MediaMetrics, and was a further refinement of the BBDO LP model. It featured the ability to analyze the degree to which vehicles excluded from the final schedule were close to justifying inclusion, a parameter which CEIR called "sensitivity".

The MIT model, designed by Ithiel DeSola Pool, then a professor at MIT, was of a completely different nature from the prior models, and was called Simulmatics. Audience data were used to create a simulated population, and the "Monte Carlo" method was used to simulate the exposure patterns among this population caused by the use of specific media vehicles.

The Monte Carlo method works as follows. Assume that a particular television program has a rating of 10.0, meaning that it reaches ten percent of the population in a single airing. In order to apply this to a simulated population, members of that population are assigned probabilities of per-telecast exposure to that vehicle.

Instead of every member of the simulated population having a .100 probability, probabilities were varied based on demographics, in line with Nielsen ratings across demographic groups for that show. For example, an adult woman 50+ in a household with many non-adults might be assigned a probability of .125 for a specific show, based on Nielsen (larger families and older people in general tend to watch more TV, and therefore have higher ratings for many shows).

In the Monte Carlo method, a random number between .000 and 1.000 is generated for every vehicle for every member of the simulated population. For example, let's say that the random number generated for the adult woman 50+ just described, for the specific

program, was .090. Since this number is below the .125 probability assigned for that simulated person for that show, that simulated person would be scored as having been exposed to that show. If the random number were .666, for example, the simulated person would be scored as not exposed to that show. In general, any random number lower than or equal to the exposure probability results in an exposure, while any random number above the probability results in a non-exposure.

If multiple insertions in the same program across weeks were to be simulated, a simulated person scored as not exposed in one week could be scored as exposed in a different week.

The advantage claimed for the Simulmatics approach was that it was more granular, down to the level of individual people in individual households, so that the exposure patterns of alternative schedules could be more definitively studied. In particular, non-linear elements of reach, frequency, frequency distribution (the tendency for a minority of heavy viewers or regular readers, etc. to receive a disproportionate share of exposures from a particular medium), audience accumulation and duplication, could be more precisely dealt with.

The Simulmatics approach prefigured the "database marketing" and "relationship marketing" ideas of the 80s, and their further evolution into the "one-on-one marketing" ideas of the 90s.

Other media optimization approaches were also studied in the early 60s. These included incremental optimization, hillclimbing, and the brute force method.

Incremental optimization requires a database of individual respondent data or a simulated population, and places emphasis on maximizing unduplicated reach. It works by selecting one vehicle at a time, removing from the analysis the people reached by the first vehicle, and then analyzing which is the best vehicle based only on the remainder of the population left unexposed by that first vehicle. The process continues, analyzing the next vehicle to buy so as to most efficiently reach the population still not reached.

The method may also be applied using the concept of effective frequency, whereby a person is not considered reached until he/she is reached a specific number of times e.g. 3 times.

Hillclimbing randomly generates a large number of schedules and picks

the best of these. Then it continues to randomly generate additional schedules until it finds one that "beats" the first selected schedule. It continues in this way until the user is satisfied that the system is not generating significantly better schedules often enough to justify continuation of the process.

A variation of hillclimbing, known as nonrandom hillclimbing, detects characteristics of winning schedules and then "climbs the hill" in the direction of these winning types of schedules, e.g. those with more daytime TV, or newspapers, or whatever type of schedule appears to be delivering the best results.

The brute force method generates all possible combinations of media vehicles and is generally impractical except where the number of vehicles can be pre-limited by a heuristic as in the later TRA optimizer.

At Kenyon & Eckhardt (now Bozell, Jacobs, Kenyon & Eckhardt), working for Norm Hecht, Harvey conceived and designed two optimization models, one which was the first local media allocation model for deciding on an automated basis how much spot TV weight to buy in each market, and the other which attempted to use the agency's advertising testing results, compiled under the leadership of Sy Lieberman and Ted Dunn, to create objective media impact weights.

Work on the second of these two models led Harvey to the conclusion that media impact weights could not be generically created so as to be applicable for all brands at all times, since the impact relationships among media would be influenced by the product category, the creative execution, the audience of the medium, the medium's environment (e.g. the mood created, the credibility of the publication, etc.), and the payoff measure which the impact weight reflected.

For example, Schwerin studies in the 50s had established that food commercials produced higher recall and persuasion when exposed in comedy shows as compared to drama shows on the radio, while proprietary research in which Harvey was involved in the 70s showed that pain reliever commercials on TV gained impact in drama shows as compared to comedy shows on TV. Clearly, impact relationships across media could be reversed for different product categories.

For another example, if the payoff measure were to be advertising recall, younger people are known to have better ability to recall, and so media with younger audiences would have an impact advantage. The resulting impact weights would be different from those that would

be assigned if sales, for example, were to be the payoff measure.

The interweaving of all of these variables to produce different sets of media impact weights for different clients under different conditions led Harvey to consider the use of direct response measures as the best media impact weights for any specific client campaign. However, direct marketing at the time was in its infancy, was assumed to be limited to direct mail, and was not used by most major advertisers.

The most important takeaway to Harvey from his work at K&E was that each ad probably had its own media impact weights by vehicle. This was a mindboggling notion of which he was intuitively convinced from his effort of poring through 1500 copy tests and trying to find a formula to derive media impact weights that could be used by all brands and ads going forward.

Ad creative was the single biggest factor that affected media impact: **the media impact scores would have to be different for each ad.**

Harvey realized that if television could be made to have an interactive element, then major advertisers would employ direct marketing measures through interactive TV, thus supplying objective measures of media impact with which to optimize schedules based on initial tests measuring incremental sales effects by direct response tied to specific vehicle codes. This was to set Harvey off in the direction of making TV interactive, so as to be more rationally optimizable.

Improved Media Optimization Models of the Later 60s and Early 70s

Harvey became Manager of the Applied Science Division of Interpublic in the mid-60s. This Division brought together media researchers and operations researchers to create systems to improve the effectiveness of media decisions for the clients of all Interpublic agencies. The systems were collectively known as Media Investment Decision Analysis Systems, or MIDAS for short. For someone with Harvey's interests, this was the place to be at the time. The leaders of the effort were Bob Coen, David Silverstone, and Larry Young, the first two media researchers and the third an operations researcher.

At Interpublic, Harvey arduously built by hand a 1000-person simulated population reflecting characteristics of demographics and media behavior drawn from Nielsen, Arbitron, Simmons, Starch, and other syndicated survey sources. Larry Young wrote an algorithm called SCANS (short for SCHEDULE ANALYSIS by Simulation) which

utilized Harvey's simulated population, and selected media so as to focus exposures within the range of frequency taken to be the most effective for the particular campaign.

This was in effect an improved version of MIT's Simulmatics model, with more exactitude in matching data to syndicated sources, and with a new emphasis on controlling frequency.

The tendency for TV advertising to "pile up" frequency on 20-40% of the population was just becoming known, and drove this new emphasis. The resulting schedules tended to utilize more dayparts of television in combination with larger lists of magazines, as compared to the schedules which had been used prior to SCANS.

While at Interpublic, Harvey wrote a document called *The Influence of Commercial Environment Upon Communication*, a major "blue book" which compiled all available real media impact data existing at that time, and drew general conclusions from these. The blue book was widely circulated for many years to all Interpublic agencies and clients.

The group also created a process called a Total Media Audit which analyzed a brand's media exposures for every medium used by the brand, at the county level, and then aggregated the data up to unduplicated TV markets and regions. Spot TV and in some cases spot radio were then recommended for use in specific markets where the audit revealed that media weight fell short of ideal as determined by the market's sales data. The approach was heuristic rather than algorithmic in its optimization of geographic delivery. Use of this tool tended to shift budgets from 80% network TV/20% spot TV to 60% network TV/40% spot TV.

An important part of the Total Media Audit was the allocation of television exposures down to the county level. The Interpublic agency McCann-Erickson had special need to deal with county data in that their client Coca-Cola had over 500 bottlers who were assessed portions of the overall advertising budget based on the estimates of exposure delivery to their respective bottler areas.

However, there were no real, current data with which to accurately allocate network and spot TV exposures down to the county level. Instead, years-old coverage data were used to estimate these allocations, which were the same for every program on a given station, even though logic indicated that the geographic pattern of a program's audience would vary as much as demographics were observed to vary program by program within station. The system which performed

these estimates was called TVCRI, for Television County Rating Indicators. Part of Harvey's job was to revise the TVCRI every time a station changed its power, tower height or location, added a retransmitter, or commissioned a special coverage study.

With sales data just beginning to become available by market through the computerization of warehouse withdrawals systems, Harvey recognized the importance of being able to accurately relate advertising delivery to sales results on precisely matched geographic areas. Interpublic's Total Media Audit appeared to be making a major contribution in this regard, lining up advertising inputs with sales outputs.

However, Harvey was troubled that his own and others' judgment estimates played such an important role in the process. It was clear that a better job could be done if real, current data could be obtained by finite geographic areas, rather than having to make estimates of such precise geographic delivery patterns.

As a result, Harvey joined the American Research Bureau, then known as ARB, and today known as Arbitron. He intended among other things to create a system of real data that would parallel the TVCRI estimating system. This became the Area of Dominant Influence or ADI. Nielsen quickly followed with the Designated Market Area or DMA system. Both the ADI and DMA divided the country into unduplicated TV markets, assigning counties to the market whose stations aggregated the highest viewing shares of that county's homes.

The name "ADI" was offered by Ace Kellner, then VP Station Sales of ARB. The area definitions had been presaged by the Marketing Area created earlier by ARB by Roger Cooper and Jim Rupp, as well as by similar definitions created by Gus Priemer, then of P&G. The ADI added every-book measurement plus the analysis of spill-in and spill-out by individual spot audiences.

This grid of unduplicated markets then became the geographic sieve for analyzing each program's audience delivery. It became possible to see that a spot on a Boston station delivered 30% of its audience into Providence, and so on. Now it was possible to compare advertising and sales data for the same pieces of geography without having to "make up" some of the underlying numbers.

CEIR then owned ARB, and Harvey persuaded CEIR to create an optimization model which started from gross rating point (GRP) goals by ADI, then backed into how many GRPs to actually buy in each

market so that the effects of spill-in and spill-out could be exploited. For example, if the goal is 100 GRP in Boston and 100 GRP in Providence, because of spill from each market to the other, buying 90 GRP in Boston and 75 GRP in Providence will actually deliver about 100 GRP in each of these markets. This obviously saved the advertiser considerable money. As a result, the ADI became "the most widely used marketing tool in the world" according to *Sales & Marketing Management* magazine.

While at Arbitron, Harvey also participated with Arbitron CEO Peter Langhoff in a study of radio impact, the development of massive customized spot TV postevaluation systems for P&G and Bristol-Myers, the development of the ARBAR system for merging audience and commercial monitoring data, the SNAP spot TV allocation/"network fill" system, and numerous other tools related to media optimization.

Working at C.E. Hooper later in the 60s, Harvey assembled NBC, Columbia Pictures, and four of the top ten agencies to create The Hooper TV Commercial Impact Index, a study of presence in room during commercials and recall of commercials, based on the robust coincidental method extended to collect media impact data. The study utilized an unprecedented sample of 250,000+ telephone interviews and was therefore the largest study of media impact in history.

In the late 60s and early 70s, Harvey was Executive Vice President of Brand Rating Research Corporation, suppliers of Brand Rating Index, the first singlesource product/media syndicated service, today mirrored by MRI and Simmons in the U.S. and by others in many other countries. These services measure media usage, demographics, geographics, psychographics, and product/brand usage, among an annual sample of people, so as to be able to cross-analyze any characteristics against one another.

In concert with the Advertising Research Foundation, the BRI sample was turned into a simulation and optimization model called first COMPASS and then later called COUSIN. 16 of the top 20 agencies met monthly to refine and utilize this model. Underpinned by the 20,000 annual BRI sample, this model was a refinement of the MIT Simulmatics and the later Interpublic SCANS models. It utilized a combination of Monte Carlo and incremental optimization approaches.

In order to improve media impact data availability to the industry, BRI CEO Norton Garfinkle and Harvey added questions to the BRI questionnaire measuring dimensions of magazine editorial environment influence upon advertising communications, and

experimented with the collection of brandswitching information as a media impact measure.

Harvey created the first online radio planning system SONAR (System for the ONline Analysis of Radio) which became part of the suite of Telmar systems. This was not a true optimization algorithm but rather a set of heuristics. SONAR became available online in 1971 via the original extremely slow modems.

With BRI consultant Sid Mehlman, formerly of Benton & Bowles, Harvey conceived the Storyfinder model in 1971. Harvey later sold this model to IMS in the 70s. The model optimizes sales presentations by media to agencies, using syndicated databases such as today's MRI and Simmons. Storyfinder has been widely emulated and has become the most used computer system by the magazine industry.

Media Optimization Systems of the 70s and 80s

In general, media optimization played a diminished role during this period in the U.S., while the 60s media optimization developments in the U.S. moved offshore and were emulated and refined in Europe and South Africa.

The big development of this period in the U.S. was the shift to supermarket scanner data. Scanner data provided the hard sales results that had previously been only a dream, and were available on a daily basis for those inclined to react that quickly. As a result of the compelling nature of this datastream, attention to other matters was occluded.

Agency research departments shrank, as did advertiser investments in consumer research by any means other than scanner data. Where consumer verbal responses were still sought in order to explain the "why" of the scanner sales data, advertisers began to fall back on small- sample focus group techniques as a replacement for full-scale quantitative surveys. These trends are still operating today, and tend to work against the success of marketers, since focus groups are a dangerous replacement for real research. But the reliance on scanner data psychologically mitigates against consideration of higher quality survey methods.

Relatively simple and less costly media optimization systems continued to be offered in the U.S. and were used on an everyday basis through online and PC-based applications offered by IMS, Telmar, and MSA. The emphasis on accuracy gradually disappeared, and users were

trained more in how to use the computer than in how to assess whether the underlying models were accurate or not.

Harvey himself took a break from focusing on optimization during this time frame. Instead his focus became new media. Through most of the 70s, he consulted on new electronic media initiatives such as QUBE, cable advertising, new forms of direct marketing (database, relationship, curriculum), and the original pioneering work in the development of audiotex (computerized telemarketing both inbound and outbound). All of these were approaches he believed would add to the ability to actually measure media impact as part of the process of marketing and advertising, and which would therefore some day allow him to double back to improved media optimization armed with real impact data.

Later in the 70s Harvey returned to optimization. Now that scanner data were available, it seemed insufficient to do media optimization merely against exposures. Why not create an optimization system which would optimize sales, rather than just optimizing the most exposures (weighted by various dimensions of value)?

The original work on the media optimization system to maximize sales was spurred by a consulting assignment from CPC International (Hellman's Mayonnaise, Skippy Peanut Butter, etc.). Harvey analyzed several years' worth of sales and audience data by market for five top brands. This led to the surprising conclusion that marketers were overspending in certain markets and underspending in others.

The markets which tended to receive excessive advertising weight were those where the brand sales per capita (BDI, or Brand Development Index) was highest. These markets were "topped-out" and further increases in advertising did not add to sales. On the other hand, markets where the BDI was rapidly increasing ("fastgrowth markets") were receiving insufficient advertising support. More weight added to the latter markets made a difference, where adding the same effort to the high BDI markets made no difference.

Harvey was allowed by CPC International to reveal these findings in his newsletter, where they were picked up by the Association of National Advertisers. ANA invited Harvey and his colleagues, then including Arch Knowlton, formerly of General Foods, to present their findings to the industry. This resulted in increased awareness of the desirability of "Momentum Marketing", or adding advertising weight to fastgrowth markets. This phenomenon added to the economic growth of growing markets and helped many of the Southern tier of U.S. cities to

consolidate their gains into major market status during this period.

The First Opti*Mark

Harvey began to hone this idea into a new optimization model called Opti*Mark. Working with Time Inc., which at the time owned warehouse withdrawal and scanner sales measurement systems under the name SAMI, Broadcast Advertiser Reports (which later became part of Arbitron), ANA, MSA and others, Harvey developed Opti*Mark as the first media optimization model to maximize sales rather than exposures. "Opti*Mark" was a contraction of the phrase "Optimized Marketing".

The first Opti*Mark system recognized the S-curve as the underlying shape of a brand's sales growth market by market. Markets at the beginning of growing a brand and those which had brought the brand to its highest potential were not fruitful places for heavy-up of advertising, while markets in between these two conditions were the best places to put incremental advertising weight, because that's where advertising made the most difference. Return On Investment (ROI) would be maximized by identifying where the next ad dollar would make the most positive difference in sales—and this is what the first Opti*Mark was designed to do.

Increases in advertising were analyzed against increases in sales to identify the markets where the change in advertising made the most positive *difference*, and then these markets were allocated increased support.

In addition, the model studied where advertising was having the most positive effect to detect if certain media types were being more prevalently used in those markets/counties, in the hopes of finding outstanding media types in terms of sales performance for a given brand at a give time. It turned out that the model did identify certain media types as more effective—and different media types were most effective for different brands at different times. The first Opti*Mark automatically identified these media types and then automatically modified the recommended media schedule to make more use of these media types for the relevant brand.

Another influence began to permeate Harvey's work during the early 80s. Harvey had much earlier observed that non-rational aspects were important in getting optimization models used—specifically their ease of use and the "sexiness" of their onscreen displays. Now around 1980 Harvey met Dave Davison, whose company Iconix was co-located with

an unknown company called Apple in Cupertino, California. Iconix was the first to base a business around the idea of creating large-screen computer displays with touchscreen control for executive "War Rooms", in which icons would be used in place of text. This idea had actually originated with Xerox in Palo Alto.

Harvey and Davison so impressed General Foods with the combination of ideas that GF acquired the first Opti*Mark model lock, stock, and barrel. The model was further developed by GF but the name "Opti*Mark" was not of interest, and so this name reverted to Harvey's use.

Meanwhile the unknown Apple company took the icon idea to its natural extension and created the Apple computer brand based upon it. Another unknown, Bill Gates, would later create Windows in emulation of the same icon idea. Dave Davison became consultant to the Secretary of Defense Cap Weinberger and helped create icon-based "Tactical Operation Centers" (TOCs) for the U.S. Army. Within the U.S. military these icon-based displays would become a standard in all branches of the armed services, including the "dashboards" of advanced military aircraft and tanks. Harvey decided that "Sexy Dashboards" would always be a feature of his optimization systems from that point on. Subsequently, the U.S. Army became a consulting client to Harvey's company, New Electronic Media Science, Inc.

In the early 80s, Vitt Media, then the leading independent media buying service, asked Harvey's company to consult on the development of the Williams Media Planning Model (WMPS), an optimization system which used a client questionnaire to specify all of the relevant goals and considerations to be used in media selection. Since similar media optimization systems existed by this point in all of the media centers of the world, in order to create relevant differentiation, Harvey spurred the creation of the Media Impact Data Bank (MIDB), a compilation of the world's resource of media impact studies. MIDB helped Vitt sell WMPS usage to its clients because the collected wisdom of the race as regards media impact was now in one place to aid clients in filling out the WMPS questionnaire with more than purely subjective "gut feel" scores.

In the mid-80s, Harvey's company consulted for R.D. Percy, whose unique passive peplemeter utilized infrared radiation to detect people's presence in the room with a playing TV set independent of the need for buttonpushing by peplemeter respondents. This provided a measure of commercial audience as opposed to program audience, which had always been used as a surrogate for commercial audience.

Percy was yet another attempt to provide definitive media impact data, the missing link in true media optimization.

In the late 80s, Arbitron commissioned Harvey's company, now greatly augmented by the presence of Len Matthews, to build utilitarian optimization systems around their ScanAmerica service, which was the first to use scanner and peplemeter measurements upon the same probability sample.

Both the Percy and ScanAmerica efforts were, like the AGB peplemeter (the first peplemeter to be used in the U.S.), unable to sustain the burn rate necessary to displace the powerful Nielsen monopoly on the measurement of television audiences in the U.S.

Meanwhile, the refinements of optimizers in the UK, Europe, Australia and elsewhere were surpassing the U.S. state of the art.

Media Optimization in the 90s

Media optimizers in the U.S. had been a flash in the pan in the 1960s, and their mainstream popularity had died out for a while after that. However, in Europe and around the world they had remained in vogue and became commonplace. By 1995, U.S. interest in media optimization resumed as a result of Procter & Gamble and Unilever, whose agencies had been using optimizers for the two companies in other countries. The immediate demand for optimizers was filled in the States by three imports from the UK, X*Pert, SuperMidas, and Figaro, and Spot On from Australia. These were daypart optimizers rather than optimizing at the program series level.^{4,5}

Although in some cases an optimizer had improved the target reach cost efficiency of a schedule by as much as 40% or more, the average worldwide experience was a reliable minimum 14% improvement when case studies were rolled up for daypart optimizers. The rare program level species tended to get significantly higher lifts than that, on the order of magnitude of about double. Still, all of these were improvements only in the cost efficiency of gaining reach against typical sex/age demos. The lack of progress in the standardization and objectification of media impact weights still persisted.

In the 90s, Harvey had a company Next Century Media which introduced the first addressable TV commercials, first set top box data to research grade, a programmatic buying/selling interface for the addressable commercials, a machine learning based personalized program recommender, and a win/win optimizer which simultaneously

optimized for both the buyer and the seller. The latter functionality was achievable because addressable commercials allow the increase in the number of target impressions per dollar, yielding incremental value, which can be shared 50/50 or any other way between buy and seller. The split of the incremental yield was left to the negotiation in each ad sale, and once negotiated, was entered as a single dashboard entry e.g. 65/35 in favor of the party with the greater leverage, etc.⁶

NCM also built the first optimizers for Nielsen Respondent Level Data (RLD). These were program series level optimizers for regular (not addressable) television and the clients were BBDO, Turner, Discovery, USA Network, and The Weather Channel.

The 20th Century ended with the media impact question unsolved. All optimizers were still limited to maximizing target reach, not campaign impact. That was considered good enough. Maximizing target reach could only help campaign impact, it was widely felt. The assumption was that media impact factors could not be so strong as to overturn reach/frequently completely – and if they were, well, heaven help us.

Media Optimization in the 21st Century

Harvey and partners, by licensing technology from NCM, created a company called TRA, which was the first company to prove that the sales effects of advertising could be measured by using set top box (and digital ad pixel) data matched with purchase records at the same-household level behind a privacy shield. This came along just in time to replace P&G's Project Apollo which proved that small "single source" panels could be created that would allow TRA-like analytics to be performed for large-penetration brands. Apollo like its predecessors AdTel, BehaviorScan, ScanAmerica was not economically sustainable by the small number of large brands that could be effectively helped at the economically feasible sample size levels, and was going out of operation as TRA was launching.

The TRA optimizer was an incremental optimizer like others Harvey had built, adding one spot at a time, in this case for the first time based on millions of homes "Respondent Level Data". Reach and duplication was measured not estimated.

For the first time, significant inroads were made against the media impact problem. This was achieved because of TRA's ability to measure ROI outcomes of different schedules. How to use this in the optimizer was the question.

The answer came in the analysis of large numbers of schedules against sales effects results. Harvey discovered that 80% of the ROI lift produced by TV advertising comes from what he called "Heavy Swing Purchasers". HSPs are defined as heavy category purchasing homes who have bought the client's brand in the lookback period (preferred implementation 3 years). Because they have previous experience with the brand, it requires far less persuasion by the ad to get them to buy it again. Targeting HSPs, because this is non-addressable television, does not mean that all you reach are HSPs; but by targeting HSPs instead of (or in addition to) a sex/age group, you also tend to get more category purchasers in general than you would in a sex/age buy.

Thus, when the TRA optimizer was used, targeting HSPs, this was akin to having media impact weights in that it tended to maximize campaign ROI impact. It differentiated program series based on their average index against HSPs, so it looked like media impact weights. HSP targeting became a proxy for ROI Optimization.

Dave Morgan created a company Simulmedia and engaged Mark Green to build its optimizer. Dave later engaged Harvey to evaluate the optimizer. Harvey studied the most recent 72 campaigns (12 months worth) and compared the cost per reach point of the Simulmedia buys to those of the relevant agency of record (AOR) for the same client. Harvey found that the Simulmedia optimizer was able to achieve much faster reach than the AORs. Harvey in presenting findings at ARF theorized that part of this was the optimizer quality itself (program level where the agencies were known to be using only daypart optimizers) plus the fact that Simulmedia uses set top box data conformed to Nielsen data. The set top box data because of its sample size (millions of homes) is far more able to more precisely see the duplication patterns among low rated programs (most programs today) than can be resolved using only small (35,000 homes) panel data.

Digital advertising brought with it a new type of optimization system called a Demand Side Platform (DSP) which optimizes digital buys. Increasingly today those are buys of digital television (OTT/CTV), both Premium (professional network TV level e.g. Hulu, CBS All Access, Crackle, Tubi, IMDb, etc.) and Non-Premium (user generated content e.g. Youtube, Facebook video, etc.).

These optimizers generally attempt to find existing client brand purchasers so as to get credit for sales that would have happened anyway, in a dissonant echo to the TRA HSP approach because the outcome for the client brand is not improved by reaching habitual

brand buyers that would have bought anyway. The optimizers of the future, in all media, will need to stay focused on outcomes involving truly ***incremental*** sales and branding effects. This will be the true gauge of the goodness of an optimization platform. The ARF (partnered with Bill Harvey Consulting, 605, and Central Control) has set up a “test rig” called RCT-21 for the purpose of establishing incrementality truth sets against which ROI measurement and optimization systems can be tested.

In 2018, Harvey’s newest company RMT published an ARF paper entitled “Crossmedia ROI Optimization Must Include Creative”⁷ which laid out some requirements for the optimizers of the future. Harking back to his earliest days in the business, Harvey’s paper emphasized the need to have impact weights reflecting the degree to which each media vehicle environment helped or hindered the sales and branding effects of the specific ad. RMT has developed a system for doing that called DriverTags™. In 2018 the ARF published a Turner/RMT paper reporting a Nielsen Catalina (now Nielsen NC Solutions) study sponsored by Turner which validated the DriverTag™ system as producing double-digit increases in sales effect via the program level media impact weights RMT calls “Ad-Program Resonance”.⁸ DriverTags™ were also independently third-party validated by Simmons in 2017⁹, 605 in 2019¹⁰, and Semasio/Reset Digital in 2020¹¹.

The DriverTag™ system can be used two ways simultaneously if one has the right media optimizer:

1. To maximize the resonance between a specific ad and a specific consumer by matching the motivations in the ad to the motivations of the consumer. RMT with partner Semasio now has this motivational data (fully privacy protected) on 276 million Americans. The intention is to use both purchaser targeting oriented to incremental sales, and motivational resonance with the specific creative, together to optimize targeting based both on product proclivity and specific ad receptivity.
2. To maximize the psychological resonance between a specific ad and a specific media vehicle context (program series, network, network/ daypart, website, etc.). This is Harvey’s solution to the media impact problem he set as a challenge to himself early in his career.

Both approaches are centric to the specific creative, as Harvey had long ago concluded would be necessary for media impact weights to be accurate.

In September 2020 at the ARF AUDIENCExSCIENCE Conference, RMT announced that by April 2021 it would make available a plug-in optimizer that anyone can use to make the DriverTag™ system easy to use within an agency, advertiser, or media planning/buying/selling system.

OptiBrain

Bill Harvey Consulting worked with Mediabrain (Mark Green, creator of the Simulmedia Optimizer, and Nick Ellis, one of the creators of X*Pert) and McKenna & Associates (Bill McKenna, a SuperMidas expert from his time as CEO of Kantar Media NA) in 2020 to create a custom optimizer for a major DTC advertiser. BHC and MckA were impressed with Mediabrain's ability to create new optimizers quickly and efficiently. Compared to software developers worked with over the years, Mediabrain is able to move more swiftly, and for the example the optimizer which RMT promised for April 2021 is available now in December 2020, and is called OptiBrain. It is not the creation of RMT but of the Mediabrain BHC MckA consortium. RMT is one optional data feed to OptiBrain.

OptiBrain is a cost-effective cloud-based SaaS user self-serve system available on an affordable subscription basis that can work in any country with any database whether panel based or set top box/ACR/digital pixel data based, whether the user has access to report level data or respondent level data.

OptiBrain can be quickly and affordably customized.

During set up with a client, OptiBrain can be seamlessly integrated into the client's relevant existing systems stack including DMP, CRM, etc.

OptiBrain comes with its own machine learning system whereby the reach/frequency estimates can learn to come closer and closer to a truth standard as specified by the user. This can be used for example to constantly train the system using big data to match the currency in the relevant country.

BHC and its partners in OptiBrain were especially motivated to include the machine learning functionality because of the World Federation of Advertisers/Association of National Advertisers "North Star" initiative, in which a blueprint implies the simultaneous use of big data and panel data, with feedback loops to provide a single view of reality. OptiBrain provides the ready technology to fill that slot in the WFA/ANA plan.

If the user is an RMT client, media impact weights are automatically applied to make OptiBrain an ROI Optimizer, not just a reach optimizer.

OptiBrain can optimize local schedules on top of national schedules, of special relevance in the U.S. with its 210 local markets. Local will gain in importance as markets come back economically from the pandemic at different rates of speed. www.optibrain.io

This brings us up to date on the history of media optimization on Earth so far. We look forward to updating this document again to include what is sure to be a flowering period in the media optimization saga.

References

¹Ephron, Erwin, "Recency Planning", *Journal of Advertising Research* 37, 4 (1997): 61–65.

²Ephron, Erwin, "Point of View: Optimizers and Recency Planning", *Journal of Advertising Research* 38, 4 (1998): 47–56.

³Lerner, David, "The Translation from Theory to Practice", Presentation before the Eastern Annual Conference of the American Association of Advertising Agencies, November 16, 1961.

⁴<https://www.adweek.com/brand-marketing/media-agencies-high-stakes-big-jackpot-27535/>

⁵<https://silo.tips/download/television-optimizers-did-they-change-the-way-we-do-business>

⁶<https://www.billharveyconsulting.com/pdf/CTAM2QuarterlyJournal.pdf>

⁷Harvey, Bill, "Crossmedia ROI Optimization Must Include Creative", Proceedings of the Advertising Research Foundation, 2018.

⁸Harvey, Bill and Shimmel, Howard, "Quantifying the ROI Impact of DriverTag™ Context Resonance", Proceedings of the Advertising Research Foundation, 2018.

⁹Pellegrini, Pat and Hutton, Graeme, "Empowering ROI by Connecting Psychographics & Programmatic", Proceedings of the Advertising Research Foundation, 2017.

¹⁰Harvey, Bill and Karam, Fadi, "Accelerating Brand Growth Using Psychological Resonance", Proceedings of the Advertising Research Foundation, 2019.

¹¹Harvey, Bill, McKenna, Bill, Cantu, Charles and Skou, Kasper, "Integration of Personal and Product Motivations Yields KPI Lifts", Proceedings of the Advertising Research Foundation, 2020.