An Operational Approach to Eliminating Backlogs in the Social Security Disability Program

Gerald Ray and Glenn Sklar

June 2019

Abstract

More than two million people file disability claims with the Social Security Administration (SSA) each year. Periodically, the numbers of both requests for hearings and appeals of hearing decisions have surged, causing backlogs and processing delays. These processing delays can have a significant impact on the health and well-being of applicants for benefits.

This paper describes some of the reasons that these backlogs occur and offers ideas, including some proven effective at the Appeals Council and hearing operation, that can help SSA eliminate these backlogs and processing delays. More specifically, this paper offers staffing, performance management, and training ideas to boost employee performance. Additionally, the paper describes how decisional support tools can be used to enhance and standardize case review processes and how data analytics can both uncover hidden problems in case adjudication and enhance case screening initiatives. The paper also offers ideas for enhancing quality assurance efforts so as to reduce rework and detect and prevent fraud. Finally, the paper offers ideas on how to structure policies to improve understanding by case adjudicators and judges and to facilitate greater opportunities for applying natural language processing and artificial intelligence to case adjudication.

Author Affiliation

Gerald Ray
Former Deputy Executive Director
Office of Appellate Operations
Social Security Administration

Glenn Sklar
Former Deputy Commissioner for Disability Adjudication and Review
Social Security Administration

The McCrery-Pomeroy SSDI Solutions Initiative

Introduction

Applicants seeking disability payments from the Social Security Disability Insurance (SSDI) program reside in every state and congressional district. While there may be varying levels of political support for the program, the authors believe that as long as the Social Security Administration (SSA) administers the program, the agency should provide timely service at all levels of the administrative process to applicants who file and appeal claims with the agency. This paper will describe effective ideas and techniques that SSA can employ to reduce claim processing times for claimants awaiting a hearing and decision on their disability claims.

The process of seeking disability payments begins when an applicant files a claim. Generally, the applicant's claim is initially adjudicated by the Disability Determination Service (DDS) responsible for the state, district, or territory in which the applicant resides. The applicant, or claimant, may request reconsideration of the initial determination in most areas of the country if dissatisfied with the determination issued by the DDS and may request a hearing if dissatisfied with the reconsideration determination.[1] The agency provides the claimant an additional administrative appellate step, the filing of a request for review of the hearing decision with the Appeals Council, if the claimant is dissatisfied with the hearing decision. Thereafter, a claimant may appeal to federal district court.

Agency processing of a request for hearing is multi-staged and usually includes the gathering of additional evidence, the scheduling of expert witnesses, the scheduling of a hearing, the holding of the hearing, the drafting of the hearing decision, and the issuance of that decision. Different staff members are actively engaged in different parts of this process, so an inventory of cases at each step in the process is necessary to keep the staff fully engaged and productive. Increasing queues at any of these steps can be considered a backlog. Similarly, backlogs can develop at the next step in the appellate process, when dissatisfied claimants appeal hearing decisions to the Appeals Council. Simply put, backlogs in the Social Security disability program are created when the available staff is insufficiently productive to keep pace with receipts and pending workload levels, which can cause claimants to wait an inordinately long time to have their cases heard and decided.

Regulatory and due process requirements build some time into the process to ensure that claimants receive timely notice of agency actions relating to their cases, ample time to submit evidence and briefs, and sufficiently advanced notice to plan for attendance at the hearing. The point where the number of disability claims awaiting adjudication exceeds a reasonable pending inventory and becomes a backlog is not precise but rather is dependent upon the number of adjudicators available to process the claims and the length of time considered reasonable for claimants to wait before being granted a hearing or receiving a decision. Some judgment related to the reasonableness of the average wait time is necessary in assessing whether a backlog exists. Currently, the agency calculates the hearing backlog to be the number of cases in excess of the number of cases it anticipates it can process within 270 days.

---

1   SSA tested the elimination of the reconsideration step in the appellate process in ten states. In those states, claimants dissatisfied with the initial DDS determination could directly file a request for hearing. The testing ended in five states in January 2019 and is scheduled to end in the remaining states in June 2020. See 20 CFR 404.906 and 416.1406, 64 FR 47218 and 83 FR 63965. Federal Register notices available at https://www.federalregister.gov/documents/1999/08/30/99-22421/modifications-to-the-disability-determination-procedures-disability-claims-process-redesign, and at https://www.federalregister.gov/documents/2018/12/12/2018-26803/modifications-to-the-disability-determination-procedures-end-of-the-single-decisionmaker-test-and. Last visited April 5, 2019.

In recent years, the number of claimants awaiting a hearing peaked at more than one million, with wait times exceeding 1,000 days in some SSA hearing offices.[2] Applicants trapped in extensive backlogs risk losing homes, vehicles, and/or private health insurance. If applicants attempt to work, they may compromise their likelihood of prevailing on their claim. If they do not work while they wait for a decision, they may face potential economic calamity. After spending a year or two in a queue for disability benefits, the job prospects of denied disability applicants may suffer.

Backlogged applications also are more expensive to adjudicate and process. With the passage of time, backlogged cases often accumulate significant amounts of additional medical evidence, including duplicative evidence, thereby complicating review of the claim prior to hearing. Waiting applicants may actually age into more generous adjudication criteria, as SSA regulations have discrete cut points at various ages (e.g., 55 years old) that enhance entitlement prospects.[3] In such instances, delays directly translate into Federal benefit awards. In short, the faster applications move at the hearing level, the better it is for all parties, including the applicant, the Federal Government and the taxpayers.

The authors believe that backlogs in the SSA disability program historically have arisen from five principle causes:

- **Economic or Demographic Factors** – surges in unanticipated receipts due to economic downturns and/or demographic shifts;
- **Litigation** – court precedent and class action litigation that resulted in the re-adjudication of claims, either by court decree or by issuance of new policy guidance to conform agency adjudication with court precedent;
- **Insufficient Staffing** – prolonged periods of inadequate staffing, sometimes caused by inadequate funding but sometimes exacerbated by delayed hiring or improper positioning of existing staff;
- **Loss of Productivity** – declines in productivity, which have had multiple interrelated causes; and
- **Programmatic changes** – legislation that expanded the potential universe of recipients or significantly altered the way in which claims are adjudicated

The causes of backlogs are varied and, to some degree, created by circumstances that are not entirely within the direct control of the agency. This paper will focus on recommended solutions the agency could adopt and implement to reduce current backlogs, mitigate their effects on the disability claimant population, and retard the creation and growth of backlogs in the future while continuing to produce high quality work products. Many of these ideas have a proven track record as they were developed and implemented successfully by the authors while running the Office of Disability Adjudication and Review and the Office of Appellate Operations within SSA. This paper will address the various ideas in five parts:

- **Staffing, Performance Management, and Training** – a description of how thoughtful allocation and management of staff and improved training techniques can enhance productivity.

---

2   The Office of the Inspector General published a graphical representation of the growth in the pending workload at the hearing level from 1993, when 357,564 claimants had pending requests for hearing, to 2016, when 1,114,079 requests for hearing were waiting in the queue. See https://oig.ssa.gov/sites/default/files/audit/full/pdf/A-05-16-50207.pdf, at page 1. Last visited January 26, 2019.

3   See 20 CFR 404.1563 and 416.963.

- **Decisional Support Tools and Data Analytics** – a description of the development of various decisional support tools and how their use can enhance productivity by standardizing case review processes, and a description of how data analytics can uncover hidden problems in case adjudication.
- **Case Screening Initiatives** – a description of how the use of differentiated case management, queueing theory, and naïve Bayes analysis can improve productivity.
- **Quality Assurance Efforts** – a description of how quality assurance and direct feedback to employees about the quality of their work can change behaviors, promote improved quality, and prevent and detect fraudulent conduct.
- **Policy and Computational Law** – a description of how policies can be designed to better promote common understanding of the policies by adjudicators, simplify case adjudication, and enhance opportunities for applying natural language processing to disability adjudication.

I.   Staffing, Performance Management, and Training

Toward More Efficient Use of Staff Time

As an independent federal agency, SSA has direct control over how it allocates its staffing and appropriated resources.[4] Generally, the agency allocates staff to operational components[5] within SSA based on budget estimates often developed with analytical models of agency workloads that include forecasts about claim filing rates, appeal rates, broad assumptions about how staff is used, how much staff time is spent on casework, how productive the staff is likely to be, and other workload trends. The estimates, calculations and models are generally reliable, but local decisions by operational executives and managers can affect the validity of the underlying assumptions of the models. This is because operational leaders often make critical decisions without consulting the modelers. For example, it is not uncommon for a line manager to redeploy a decision writer to assist with clerical case screening, to draft a response to a policy proposal, or to engage in training and mentoring. Over time and across many offices, these *ad hoc* decisions can dramatically alter operational output. Additionally, the duties of the staff morph over time as technologies are introduced into the workplace.

Budget and workload analysis modelers could incorporate these *ad hoc* changes to staffing availability if supplied with more granular information to include in their models; however, the necessary data is not always captured or made available to the modelers. Moreover, the agency has not strategically facilitated the capture of some of the information that would be helpful to workload managers as well as budget analysts. For example, the Office of Systems, the SSA component responsible for developing and

---

4   An organizational chart of the Social Security Administration can be found at www.ssa.gov/org/ssachart.pdf. Last visited April 3, 2019.
5   Operational components involved in direct casework related to the Social Security disability program include the Office of Operations, which is responsible for taking and developing disability claims, effectuating payment of benefits to claimants and their eligible dependents as well as other auxiliary beneficiaries, overseeing the state Disability Determination Services, and operating the Federal Disability Determination Service; the Office of Hearing Operations, which processes requests for hearings; and the Office of Appellate Operations, which processes appeals of hearing decisions and conducts own motion review of unappealed hearing decisions. The Office of the General Counsel also has an operational arm, the Office of Program Law, which assists in the defense in federal courts of agency decisions related to disability claims.

maintaining agency computer systems, explicitly rejected the notion of incorporating information that could be used in performance ratings into the electronic case management systems used in the Office of Hearings Operations (OHO), the SSA component responsible for processing requests for hearing, and in the Office of Appellate Operations (OAO), the SSA component responsible for processing appeals of hearing decisions.[6] While that decision may have been a sound one from the perspective that performance standards are sometimes renegotiated, SSA missed the opportunity to capture granular information about the activities of its staff. Capturing downtime and casework time opens additional opportunities to identify productivity problems, make essential staffing decisions, and drive productivity improvements.

OAO had a long history of capturing information about the time employees spend engaged in activities other than their regularly assigned duties. For example, OAO captures information on training and mentoring activities and time spent serving as union representatives. For many years, this information was captured by OAO employees, and reviewed by their supervisors, in hand-written reports. Realizing the value of this data, OAO recently automated this process in its Job Information Metrics (JIM) tool to ensure that all work hours, assignments, training, and casework are captured.[7] JIM captures granular information about the activities of the staff that helps OAO managers formulate accurate workload projections, and JIM's data underlies all OAO budget formulations and staffing requests. In Fiscal Year (FY) 2018, the Chief Judge in OHO expressed interest in revising the OHO version of the JIM tool to capture information similar to that OAO captures about OHO staff. The rest of the agency still lags behind in capturing this type of information despite backlogs in other case work such as post-entitlement work performed in the Program Service Centers.

The authors suggest that the agency periodically compare the duties of existing staff with their position descriptions to ensure that all employees are performing the core duties of their position description and to incorporate the impact of new technologies on employee work expectations.

The authors also recommend that the agency specifically track the time employees spend temporarily not performing their normal duties, particularly for those employees involved in casework. Specifically, the authors recommend that the agency develop JIM-type tools for all operational components to capture time in casework.

The authors also recommend the agency use this more granular casework information to build new workload forecasting models, to revise existing budget models, and to reallocate staff where appropriate.

---

6   OAO sits at the end of the administrative process that adjudicates disability claims. It provides staff support for the Administrative Appeals Judges of the Appeals Council, a body which operates on a direct delegation of authority from the Commissioner of SSA to review hearing decisions and assist the Commissioner in overseeing the disability program.

7   OAO staff worked with information technology colleagues in OHO to develop a local solution that automated this manual reporting by building an OAO-centric version of the JIM tool in use at the time by OHO. OAO's version of the JIM tool included specific features to capture the actual time employees spend in casework for OAO; incorporate information from its electronic case management system, the Appellate Review Processing System (ARPS), and case dispositions produced by each staff member; and then calculate the performance for each individual consistent with OAO's productivity performance expectations.

Productivity Metrics Can Help Drive Productivity

The authors have demonstrated that numeric-based performance standards enhance productivity. While challenging to establish at the outset, numeric-based performance standards can have a significant impact on service delivery, and the impact can be quantified with great accuracy. For example, OAO experienced 22 consecutive quarters of year-over-year, same quarter improvement in staff performance after implementing productivity metrics, establishing branch performance goals and revising staffing allocations.[8] In FY 2009, the 945 total OAO staff members[9] completed 89,066 dispositive request for review case actions, or approximately 94.25 per staff member.[10] In FY 2013, 1,210 staff members produced 176,251 dispositive request for review case actions, or 145.66 per staff member. In FY 2017, OAO's 1,000 staff members produced 160,776 dispositive request for review case actions, an average of more than 160.75 per staff member.[11]

In 1990, the agency suspended the use of numeric indicators in staff management, ostensibly because some irregularities and misguided management efforts had been discovered. Not surprisingly, productivity fell significantly, at least in components that continued to measure such things. OAO was one of those components that continued to monitor productivity. The elimination of productivity metrics resulted in an immediate decline in analyst productivity, from 1.8 cases per day to 1.4 cases per day, beginning with the month after numeric performance standards were suspended. Using the data OAO collected on performance, one of the authors helped persuade then-Commissioner Gwendolyn King to allow the reinstatement of numeric performance standards in OAO in 1992. Once reinstated, performance again increased, to an average of 2.0 cases per day. The decision to allow numeric based productivity standards was again reversed two years later, and OAO again experienced another significant decline in performance, although productivity only fell to about 1.75 cases per day at that time.[12]

In 2006, one of the authors developed an iterative methodology for establishing productivity metrics that provides weighted time values for various types of work products.[13] The process accounts for varying degrees of complexity in the work performed.[14] Using data from the prior case control system, he isolated

---

8   The year-over-year, same quarter improvements ended because of a government shutdown in the first quarter of FY 2014, precipitated by the failure of elected officials to timely enact an appropriations bill or continuing resolution to fund continuing operations.

9   Note that not all of the staff in OAO is directly involved in the production of request for review case actions, but all staff members were included in our assessment of the productivity increase. OAO also processes civil actions, performs quality assurance reviews, processes attorney fee requests, answers Congressional and public inquiries, and performs many other tasks. OAO realized service improvements in these other areas as well.

10 Figures cited in this paragraph were obtained from Executive Director's Broadcasts. The Executive Director's Broadcasts are published by the Social Security Administration, Office of Appellate Operations, and the information cited can be found in the Executive Director's Broadcasts dated October 19, 2009, October 7, 2011, October 24, 2014, and October 20, 2017. Copies available upon request at the Office of the Executive Director of OAO. Disposition data is also available at: https://www.ssa.gov/appeals/DataSets/AC02_AC_GrantReview_All_Dispositions.xml. Last visited April 5, 2019.

11 This figure represents a 70.5 percent improvement in productivity per staff member during the time Patricia A. Jonas served as Executive Director and one of the authors served as Deputy Executive Director of OAO.

12 Numbers cited in this paragraph are based on the recall of one of the authors, who reviewed monthly OAO reports while serving in OAO during the period in question. The reports included a chart reflecting monthly changes in productivity, and the author recalls the stark changes in performance when numeric performance plans were suspended and reinstated.

13 If work products are fungible, the calculation would simply be to divide the total number of hours of work by the number of products produced to derive the number of hours per work product.

14 This methodology was recently converted into a computer program by SSA's Analytics Center of Excellence for potential use by components outside the hearings and appeals process.

12 types of case dispositions issued by the Appeals Council and calculated time values for processing each of these types of cases based on their complexity.[15] With this information, OAO was able to establish numeric-based performance expectations in 2006 that resulted in higher staff productivity. The author used the same basic technique to develop numeric based performance metrics for decision writers at the hearing level, which were implemented by OHO in 2016 and 2017.

The productivity performance metrics developed using this technique are sufficient to set a baseline for satisfactory performance and a baseline for outstanding performance. The lower baseline serves to ensure that all employees in the job are producing a fair share of work. A slightly higher baseline bar can be set for journeyman employees, who presumably have demonstrated above-average proficiency in order to be promoted to their journeyman positions. The upper baseline sets a standard by which management can reward high performing individuals.

Although somewhat counterintuitive, OAO demonstrated that implementation of the lower baseline typically results in improved overall performance, as low producers are compelled to become more productive or see their ratings decline, putting their job security at risk. OAO also observed an uptick in performance from employees near the lower baseline, who apparently did not want to be associated with low performance, so they stepped up their performance and began performing at a rate closer to the average. Little change was noted in high performers, who were performing above the baseline for outstanding productivity before that baseline was established by the performance metrics. OAO found that a lowering of the productivity expectation for outstanding performance by 7.5 percent yielded a 12 percent improvement in overall performance, as employees who were just below the outstanding level stepped up their performance to achieve the outstanding level.

The calculations are designed to determine the average performance in each of the work output categories identified. Applying the average weighting for each case disposition for each employee yields a distribution of employee performance around that average. It is important that managers and executives understand this concept, as some may be tempted to set the bar for minimal performance at or near the calculated average. Raising the bar too high increases the likelihood that more individuals will fall below the higher acceptable levels of performance and will require more performance management activities by managers. If too many employees are put on management performance plans based on standards that are too difficult for average employees to achieve, morale and quite possibly employee retention and performance might actually decline.

Instead, managers should set the minimal performance expectations at a level at which no more than 10 to 15 percent of the employees initially are not meeting. The managers should then actively manage the underperforming employees, offering advice, guidance, counseling, and training to assist them in achieving successful performance and taking adverse performance-based action only if those efforts fail. As their performance rises, so will the performance of the entire operation.

---

15 The methodology could be improved with more granular data. Currently available agency data from case control systems track case activity and movement by date rather than by hour and minute.

In furtherance of performance management efforts, OAO organized a team of managers to monitor the performance data and intercede with line managers when performance fell below expectations at the branch or individual level. The team assists with the development of performance improvement plans, working with human relations staff and line managers to ensure intervention occurs as soon as a drop-off in performance is noted. The intent of the intervention was to bring employee performance back up to speed rather than to performance manage them out of their jobs. This approach is good for both the employees, whose performance typically improves enough for them to retain their jobs, and the agency, which invests a lot of time, effort, and money in hiring, training, and developing its staff.

The performance evaluation system SSA put in place in the early 2000s reopened the door to numeric-based productivity standards. Despite the opportunity this presented, most agency components have not implemented truly objective performance measures and continue to rely on performance plans that can be quite subjective in nature. It is small wonder that on the Federal Employment Viewpoint Surveys relatively few SSA employees agree that differences in performance are recognized in a meaningful way (only 32.5 percent agree or strongly agree in 2018)[16] and that steps are taken to deal with poor performers who cannot or will not improve (only 23.3percent agree and 7.6 percent strongly). The authors believe that objective, numeric-based performance standards could improve agency operational management and should be implemented more broadly within the agency.

It should be noted, however, that Administrative Law Judges (ALJs) are not rated under performance plans and thus cannot be placed under direct productivity performance expectations. Nonetheless, in 2015, the theory of whether an agency may establish reasonable production goals for ALJs was tested in the seminal case of *Shapiro v. Social Security Administration*, before the United States Court of Appeals, Federal Circuit.[17] The petitioner, ALJ Mark Shapiro, contested his removal from the ALJ corps due to "unacceptable performance." The court noted, however, that ALJ Shapiro had decided "drastically fewer cases as opposed to his peers," notwithstanding "extraordinary efforts" to assist, mentor, and train him. Indeed, at one point, ALJ Shapiro had "over 70 percent of the 1000-day-old cases in the New York Office." In a unanimous decision, the Federal Circuit upheld the ALJ's removal based on unacceptable performance. This watershed case provided a framework for ensuring that all SSA ALJs owe disability applicants a reasonable level of service.

The authors recommend the agency make the objective measurement of performance, including productivity, a central theme of the performance plans they issue to operational personnel.

The authors recommend that the agency make clear distinctions between levels of performance, recognize high performance where it exists, and actively manage underperforming staff.

The authors recommend that SSA data scientists assist all operational components in developing numeric-based productivity expectations for case workers.

The authors also recommend that managers and executives receive training on where to set performance levels and how and when to best intercede when performance is not at satisfactory levels.

---

16   Only 30 percent of employees rated the agency highly in this category in 2018, down from 32 percent in 2011. See https://www.ssa.gov/careers/files/11evsssa.pdf and https://www.ssa.gov/careers/files/2018FEVSresults.pdf. Last visited January 31, 2019.

17   *Shapiro v. Social Security Administration*, 800 F.3d 1332 (Fed. Cir., 2015).

Encouraging Higher Performance through Office, Branch, or Unit Goals

Some observers have been critical of performance standards and goals, suggesting that they promote behaviors that achieve higher productivity at the expense of work quality. This refrain is particularly popular among low producers and is reflected in the angst of managers who are unsure of how to improve productivity without impairing work quality. Analysis of data relating to both the quantity and quality of work produced can help determine the accuracy of such assertions.[18]

In early 2011, the authors obtained data demonstrating that ALJs moving cases through the system too rapidly eroded quality. The authors were concerned that extreme productivity might present opportunities for fraud and potentially erode program support. Feeding "super-producers" endless blocks of transfer cases from other offices also undermined the notion of random case assignment. As a result of this analysis, the authors capped the number of cases that any ALJ could clear in a year. Several "super-producers" and local managers were concerned that this cap might impact public service somewhat. Notwithstanding these objections and small loss of productivity, the authors capped the maximum processing number at 960 cases annually, the level above which the data showed significant erosion of quality.[19]

As a general proposition, the data the authors reviewed showed little change in the quality of work when the quantity increased, except at extreme productivity levels. In fact, the data reflected that most higher producers had a smaller percentage of their work returned by quality assurance reviewers than lower producers. The authors drew the conclusion that having a better understanding of how to do something often is not only reflected in improved work quality, but also in reduced rework, which aids workers in performing the work more rapidly.

Thus, to encourage higher performance among those whose performance exceeds the minimal expectation, OAO implemented group goals and found that the goals contributed to higher performance. OAO's quarterly branch goals are based on the number of analysts in the branch, their expected hours in casework during the quarter, and the calculated average expected performance for that number of casework hours. Reductions to the calculated goal are provided to branches where the branch chief has employees on performance improvement plans because it is assumed that resources will be expended in attempting to improve their performance. Branches with employees who are not performing at the minimal level do not receive the discount if the manager does not actively manage the situation.[20] This acts as an incentive for the branch manager to do his or her job. Similar goals also could be developed and implemented for hearing offices.

---

18 The development of case analysis tools, discussed in Part II of this paper, provided the authors with the type of robust data necessary to compare the quality of work with the quantity of work produced at an individual level.

19 See Statement for the Record of Glenn Sklar, Deputy Commissioner for Disability Adjudication and Review, Social Security Administration, before the House Committee on Oversight and Government Reform, Subcommittee on Energy Policy, Health Care and Entitlements (June 27, 2013) (noting new management instruction limiting assignment of new cases to 960 cases per ALJ annually).

20 The negotiated bargaining agreements between management and the unions in SSA provide that underperforming workers must be given an opportunity to improve their performance before management takes an adverse job action. In the experience of the authors, performance typically improves if the manager documents the opportunity to improve performance and develops a performance assistance plan for the underperforming employee.

The authors believe that the old maxim "what gets measured gets done" reflects why productivity metrics and goals can be valuable in driving performance. Measuring performance keeps employees focused and provides management with key performance indicators that can be used to positively intervene in work processes to enhance performance and improve results.

The authors recommend that detailed per capita-based branch, unit, or office goals be developed for each operational component and that managers be tasked with monitoring and trained on how best to monitor data related to the success of their components in achieving these goals.

Toward More Efficient Distribution of Staff

SSA's Senior Executive Service (SES) corps is relatively small when compared to other federal agencies. The hearings and appeals components in SSA, OHO, and OAO have a handful of SES executives directly responsible for operational results. These executives have a span of control far greater than most agency executives, particularly those in smaller support components that have a much more limited scope of responsibility. Given the high public and political visibility of this part of SSA's disability program, it might be wise for SSA to allocate additional SES positions to these two components.

OHO's workforce is scattered across the country in more than 165 offices of varying sizes in regions of varying sizes. Maintaining staff in some of these parts of the country is quite difficult. Certain locations experience regular attrition of ALJs, who often prefer to live in larger cities or in areas with better climates, leaving some offices chronically understaffed. Additionally, receipts vary over time by location, so the numbers of ALJs needed in a particular location can vary. The agency typically addresses the staffing imbalances with workload transfers and temporary ALJ travel to underserved locations. OHO also maintains first-line staff supervisors in each of the offices as well as managing Hearing Office Chief Administrative Law Judges (HOCALJS), who spend part of their work time on non-case related activities. Many of these non-case related functions are ministerial in nature and can be handled by others. Some of the supervisory duties could be divested to regional centers, and the numbers of managers could be reduced if OHO reduced the dispersion of staff by consolidating offices. This would free more employees to perform direct casework. Reducing the number of HOCALJs would be a particularly effective force multiplier, given the difficulty in hiring new ALJs and the backlog of cases awaiting a hearing by an ALJ.

One of the authors launched centralized case pulling and decision writing centers, as well as national hearing centers, in 2010. Consolidating ALJ hearing locations could accelerate this trend and would reduce the need for supervisory ALJs. To consolidate ALJs, OHO would first need final regulatory authority allowing SSA to determine the manner in which hearings are conducted, whether in person or via videoconference.[21] Currently, the vast majority of claimants do not object to video hearings. With

---

21   SSA published a notice of proposed rulemaking on November 15, 2018, proposing to end provisions related to the manner for the appearance of parties and witnesses at a hearing by removing the ability of the parties to opt out of video hearings. See 83 FR 57368. It remains to be seen whether the agency will press forward with final regulations. In a letter to the agency dated January 31, 2019, a number of Members of Congress objected to the notion, mostly on the grounds that in person hearings are necessary to evaluate the credibility of witnesses and claimants. Available at https://waysandmeans.house.gov/sites/democrats.waysandmeans.house.gov/files/documents/Signed%20Letter%20to%20SSA%20Regarding%20Proposed%20Video%20Hearings.pdf. Last visited April 4, 2019.

such authority, SSA could begin equipping local offices, including district offices, with hearing rooms and videoconferencing equipment for claimants to use so they would not have to travel far to hearings. Then, SSA could begin to consolidate offices, holding hearings and issuing decisions from regional centers. In theory, this also could reduce the numbers of line managers, including management ALJs necessary for efficient functioning of OHO. The consolidation should reduce facilities costs, increase staff time availability for case work, improve case assignment and communications, and facilitate training initiatives.

If SSA elects to consolidate its staff in this way, OHO might also consider creating a divisional structure consistent with the work of anthropologist Robin Dunbar. He studied social relationships and concluded that neocortex size cognitively limits the number of stable social relationships humans can maintain to about 150 people, a number known as Dunbar's number.[22] OAO's executives considered Dunbar's number when they set up the divisional structure of OAO in late 2008. OAO set up nine operational divisions, each having no more than 100 people per division. Borrowing concepts from networking theory about connectedness and betweenness,[23] OAO also sought to place key opinion leaders at the center of each division. Although each division had a hierarchical structure, led by a Division Director, who had an assistant and multiple branch chiefs as direct reports, OAO leadership instituted a separate networked structure. They created the Division Chief Administrative Appeals Judge (DCAAJ) position. Each DCAAJ only supervised only a few Administrative Appeals Judges (AAJs), and they were not involved in the active management of most divisional personnel; however, they were involved in training and in broad management decisions, and they helped establish the culture and direction of each division. The combined hierarchical and networked structure of OAO enhanced the connectedness of the executives with the operational staff, and the betweenness of the DCAAJs created an additional channel for ideas to flow from the staff to the front office, helping executives more quickly identify and resolve operational problems.

The authors recommend the agency allocate additional SES positions to OHO and OAO to reduce the span of control and improve the communications and management of these direct service operational components.

The authors recommend SSA pursue regulatory change to eliminate the opt-out provision of the manner of hearing regulation related to hearings held by videoconference.

The authors recommend OHO consolidate its staff into fewer, larger regional centers.

The authors recommend OHO reduce the number of supervisory ALJs, particularly HOCALJs, by having the strongest HOCALJs cover multiple offices. These HOCALJS would have a lighter case load and would focus primarily on core management issues.

The authors recommend OHO consider standardizing the size of its regions and consider creating a divisional structure within the remaining offices to promote the formation of stronger social and leadership networks within the regions.

---

22 See Robin I. Dunbar, "Neocortex Size as a Constraint on Group Size in Primates," The Journal of Human Evolution, Volume 22, Issue 6, June 1992, pages 469-493. Available at http://www.uvm.edu/pdodds/files/papers/others/1993/dunbar1993a.pdf. Last visited January 31, 2019.

23 In this context, connectedness refers to the number of individuals with whom one regularly interacts and betweenness is an indicator of centrality of those interactions, or how well individuals are connected through primary, secondary, and tertiary connections.

Employee Training

Training and feedback are of paramount importance in creating a high-performance workforce. Improved understanding and expertise make it easier to perform work, which leads to better quality and reduced errors and rework, thereby improving productivity. In focusing on training as a key to improved performance, OAO executives noted with interest the ideas propounded by author Daniel K. Pink, who described specific intrinsic motivators that include responsibility over challenging work, mastery of subject matter related to the work, and the ability to work autonomously.[24]

OAO executives turned to social science to develop a central management philosophy for all OAO managers. They reviewed a variety of materials relating to management theory and human motivation, including Frederick Taylor's *The Principles of Scientific Management*,[25] W. Edward Deming's 14 principles of management,[26] and Douglas McGregor's theories X and Y and his other ideas about how a manager's attitude impact employee's motivation.[27] Of particular interest was an effort by the Corporate Leadership Council (CLC) to gauge the improvements in performance that could be gained by a wide range of management actions.[28]

OAO executives developed an OAO management training course explaining what the research demonstrated were the best techniques to motivate employees. OAO executives also encouraged managers to try new techniques. If these did not work out as envisioned, managers were expected to try other things to resolve problems. Executives made clear their view that the only real negative was inactive management; that is, not trying to solve problems at all. These general efforts led to a high level of engagement by managers, judges, and employees alike, and the authors believe this effort contributed to the significant performance improvements OAO experienced. This approach also helped successfully train a new generation of leaders in OAO.

OAO executives also tasked several AAJs with developing training agendas, materials, and techniques for staff in OAO and the Appeals Council. They began by developing an understanding of how adults learn.[29] They discovered that adults retain information and understand work processes better when training includes a contextual framework. They also considered motivational techniques, particularly as they relate to learning. They identified new highly interactive training methodologies designed to teach adjudicators how to apply the law and regulations in the evaluation of disability claims. The new approach included

24   See generally, Daniel K. Pink, Drive: The Surprising Truth about What Motivates Us, Riverhead Books, Penguin Random House, L.L.C. New York, New York, 2009.

25   See generally, Frederick Winslow Taylor, The Principles of Scientific Management, Harper and Brothers Publishers, New York, New York and London, United Kingdom, 1919.

26   See generally, William Edwards Deming, Quality, Productivity, and Competitive Position, MIT Center for Advanced Engineering, 1982, later renamed and republished as Out of the Crisis, The MIT Press, Boston, Massachusetts, 1986. ISBN-10: 0262535947

27   See generally, Douglas McGregor, "The Human Side of Enterprise," available at www.kean.edu/~lelovitz/docs/EDD6005/humansideofenterprise.pdf. Last visited December 21, 2018. First published in Adventure in Thought and Actions, MIT School of Industrial Management, Cambridge, MA 1957. Copywrite by MIT, MIT Press, Cambridge, MA, 1966.

28   See www.loc.gov/extranet/cld/workforce-performance/PerformanceDrivers.pdf. Last visited December 21, 2018.
      See also https://marble-arch-online-courses.s3.amazonaws.com/CLC_Building_the_High_Performance_Workforce_A_Quantitative_Analysis_of_the_Effectiveness_of_Performance_Management_Strategies1.pdf. Last visited December 21, 2018.

29   See generally, The Neuroscience of Adult Learning: New Directions for Adult and Continuing Education, Number 110, Summer 2006, edited by Sandra Johnson & Kathleen Taylor. Published by Wiley Subscription Services, Inc., San Francisco, CA.

casework, group discussions, and exercises. This approach was a sea change from earlier training efforts, which largely were lectured based and which primarily focused on describing what the law and regulations said as opposed to how they might be applied.[30]

OAO jettisoned lecture-based new-hire training that previously had focused on teaching rules, regulations, and requirements and replaced it with more than 4,000 pages of new training materials designed around case studies, group exercises, and discussions. OAO developed scaled response surveys to measure student satisfaction and understanding of the information they were being taught. Occasionally, materials and issues were taught a second time during training class, sometimes with different instructors, to ensure the trainees learned the materials presented. The new approach made it possible to reduce the time new-hires spent in the classroom from eight weeks to six. OAO also tracked post-training performance by training class, using the data to tweak the training provided to subsequent classes. Overall, by tracking post-training performance, OAO found that the new training process dramatically reduced the new-hire learning curve, the time it took for new hires to become fully productive and demonstrate good quality, from 18 months to five months.[31]

Following initial training, OAO continues to analyze work quality. When errors are detected or opportunities are identified for improving decision quality and performance, additional training is provided. Supplemental training is sometimes delivered to groups of employees, but the primary focus of the all training is to teach and improve the performance of each individual.

The Council also developed multi-tiered training modules designed to address common errors identified in remand orders or through focused reviews and data analysis. Training specific to the issues is then provided, sometimes through in-person training but also to individual adjudicators through multi-tiered training modules available through the agency's intranet. The multi-tiered training modules are delivered through an electronic medium known as "How MI (Management Information) Doing." These modules allow employees to self-train on issues that are causing their work to be remanded. The modules are tied directly to the issues generating the remands. The modules typically contain a desk guide describing common problems related to specific areas of the law and how to avoid them; a second tier that ties into the initial training modules; and a third tier that provides the underlying legal, regulatory, and sub-regulatory language pertinent to the issue.

How MI Doing encourages employees to learn how to avoid future errors. This interactive feedback has proved quite valuable to employees and has aided in the measurable improvement now obvious in the quality of hearing decisions. OAO's training efforts twice earned the agency the prestigious Deming Award for Training Excellence from the Graduate School USA.[32]

---

30   See generally Beth Crandall, Gary A. Klein, and Robert R. Hoffman, Working Minds: A Practitioner's Guide to Cognitive Task Analysis, The MIT Press, Cambridge, Massachusetts, 2006.
      See also Judy Willis, "Research-Based Strategies to Ignite Student Learning: Insights from a Neurologist and Classroom Teacher," Association for Supervision and Curriculum Development, Alexandria, VA, 2006.

31   Gerald K. Ray and Jeffrey S. Lubbers (September 2015), "A Government Success Story: How Data Analysis by the Social Security Appeals Council (with a Push from the Administrative Conference of the United States) Is Transforming Social Security Disability Adjudication." (W.D. Hermandorfer, Ed.) The George Washington Law Review, 83(4/5), 1575-1608.

32   A list of recent Deming Award winners can be found at the Graduate School USA website at: https://www.graduateschool.edu/content/deming. Last visited December 23, 2018.

The authors recommend the agency train all managers and executives on basic concepts of management theory.

The authors also recommend the agency train managers and executives specifically on proven motivational techniques.

The authors recommend that SSA gradually replace all SSA lecture-based training for new employees with interactive training that includes exercises about how to apply the concepts taught during the training.

The authors recommend the agency build How MI Doing tools for all employees involved in operational case work.

II. The Game Changing Impact of Decisional Assistance Tools and Data Analytics

Case Analysis Tools and Decisional Assistance Tools

Over the last decade, the agency has successfully integrated several decision support tools to help adjudicators "think through" their work products. These tools allow adjudicators to memorialize their key decision points in each case. As the agency moves to more rules-based decisional support systems, there will be significant benefits for fairness, accuracy, and in-line data capture. These decision support tools are quite different than the simple case processing tools of the past that captured case movement from point A to point B over a specified interval of time.

In the early 2000s, SSA implemented an electronic case management system designed to index, house, and maintain electronic evidence for disability claims. Documents were stored as scanned images in tagged image file format (TIFF) so people accessing the electronic record could not change them. Later, the system evolved to include electronic filing of claims, evidence, and requests for hearing. Adjudicators used a Document Management Architecture (DMA) viewer to review the evidence. In 2005, the agency also implemented a new case processing system for the hearing operation known as the Case Processing Management System (CPMS). In the early 2000s, several forward-thinking ALJs developed the Findings Integrated Template (FIT), a decisional template intended to guide ALJs through the decisional process. FIT was a document-generating system that allowed the user to select the type of document they were preparing.

Templates for various types of decisions and dismissal orders included standardized language relating to pertinent statutory and regulatory authorities as well as standardized language describing the sequential disability evaluation process. FIT also included reminders to include rationale in support of conclusions and findings reached and included suggested but highly customizable findings and decisional language. FIT was not integrated into other case analysis or decisional assistance tools but was widely adopted by the ALJ corps and quite helpful in improving the consistency of hearing decisions.

Meanwhile at the Appeals Council, enterprising AAJs, led by one of the authors of this paper, also sought to capture structured data about the quality of the adjudicative work done at both the Council and hearing levels with the specific intent of analyzing data captured by analysts and adjudicators. To facilitate this effort, they mapped the policy compliant pathing to each of the approximately 2,000 types of decisions that can be made in disability claims.[33] They then crafted a series of questions designed to guide analysts and adjudicators through a structured approach for determining whether a basis exists under the law and regulations for the Council to grant a claimant's request for review of a hearing decision.[34] One of the authors then worked with computer programmers in the Office of Systems to incorporate these questions into a case analysis tool that was integrated into a new electronic case management system, the Appeals Review Processing System (ARPS), then being developed for the Appeals Council. The electronic case management system, with its integrated case analysis tool known as the Appeals Council Analysis Tool (ACAT), became operational in March 2008.[35] During the next two years, OAO staff and AAJs worked with computer programmers to update ACAT to capture structured data about cases remanded from the federal courts and unappealed decisions reviewed under the own motion authority of the Appeals Council.[36]

Concurrent with the development of ACAT, both of the authors contributed to the development of an electronic case analysis tool (eCAT) designed for use at the initial and reconsideration levels of the disability program. Like ACAT, the eCAT was designed for adjudicators to follow policy-compliant pathing during their analysis of disability issues, and, like ACAT, it captured structured data about that process. Initially eCAT was piloted in several states before implementation nation-wide a few years later.

Shortly after ACAT was developed, one of the authors tasked several ALJs, including some of those who developed FIT, to work with Office of Systems personnel to develop an electronic Bench Book (eBB). This decisional assistance tool also followed policy-compliant pathing and was designed to help guide ALJs through the hearing process as well as the decisional process. Dozens of other ALJ volunteers participated in its development from 2010 to 2014. At its peak, several hundred ALJs used this tool to prepare and/or adjudicate their cases. Although the initial iteration of eBB required the ALJs to spend a significant amount of time capturing data, the long-term vision was to slowly integrate other hearing office support personnel, including case technicians, screening personnel, and, most importantly, decision writers into the earliest stages of case preparation. That vision has not been fully realized.

The theory was that a unified electronic platform would enable lower-cost personnel to handle much of the preparation work and data capture for the ALJ corps. This value proposition was further enhanced when support personnel began to telework aggressively in 2016. Exactly how the office can function efficiently with so many players outside the office, without using a common web-based tool to support

---

33   Gerald K. Ray and Jeffrey S. Lubbers (September 2015), "A Government Success Story: How Data Analysis by the Social Security Appeals Council (with a Push from the Administrative Conference of the United States) Is Transforming Social Security Disability Adjudication." (W.D. Hermandorfer, Ed.) The George Washington Law Review, 83(4/5), 1575-1608.

34   See Office of the Inspector General, Social Security Administration, A-12-13-13039, Request for Review Workloads at the Appeals Council, 7-10, 15 (2014). Available at https://oig.ssa.gov/sites/default/files/audit/full/pdf/A-12-13-13039.pdf. Last visited January 31, 2019.

35   Ibid.

36   The Executive Director's Broadcasts are published by the Social Security Administration, Office of Appellate Operations, and the information cited can be found in the Executive Director's Broadcasts article entitled ARPS Took OAO into New Era, published on January 6, 2017, at page 3. Copies available upon request at the Office of the Executive Director of OAO.

case adjudication, remains murky. The authors are concerned that no common web-based platform exists for case preparation at this time, and we recommend that eBB access should be overhauled to expand its use to other hearing office support personnel. OHO also should expand the utility of the eBB so selections made in the eBB can be directly reflected in standardized language in decisional documents through a linked document generating system.

To gain wider acceptance and use of eBB, the hearing operation could alter its business process to have decision writers review and analyze cases for the ALJs, assisting the ALJs in identifying issues that require clarification at the hearing or additional development before the hearing. The ALJs could use the eBB to take notes, develop their questions for the hearing, and to draft their decisional instructions. This more effective use of ALJ time could yield significant productivity and quality improvements, perhaps reducing the numbers of ALJs needed to timely process the workloads. To maintain the confidentiality of the deliberative process, information and data related to that process could remain hidden from data analysis and outside review through access controls to that information and data. The eBB could be a useful tool, not only to help ALJs navigate the complexities of disability adjudication, but also to provide an additional data source about program adjudication that could be used to improve training and develop more intelligent decisional tools.

The authors recommend continued development of case analysis and decisional support tools. Increased integration of the tools with the document-generating system and propagation of information from these tools, case control systems, vocational resources, and other data sources could greatly expedite decision drafting while also improving the quality and consistency of agency decisions.

The authors recommend that the agency modify the eBB to better facilitate use by other hearing office staff, which OHO should task with performing most of the case work-up and data capture.

The authors recommend that the agency revise the hearing office business process to enable decision writers to review cases before assignment to an ALJ, so they can provide greater assistance to the ALJs by describing the evidence, identifying issues that require further development before the hearing, and identifying issues that may require clarification at the hearing. The agency could look to the operation of the Appeals Council for a model of how to implement this idea.

Data Analytics

The relatively recent availability of low-cost data storage coupled with massively expanded computational power has created tremendous opportunities to use data analytics to address backlogs. While serving as executives at SSA, the authors identified a range of areas where data analytics was helpful in improving business processes. Additionally, the authors recognized the importance of having staff with the necessary skill sets to conduct high quality analysis on data sets. The authors actively sought out employees with multidisciplinary expertise, including mathematicians, computer scientists, economists, and operational research specialists.

When ACAT was developed, OAO executives sought, and the Office of Systems granted, direct access to the underlying data of ACAT as well as the data captured in the hearing level case management system, CPMS, and in the Appeals Council's case management system, ARPS. At the time, few, if any, operational executives had access to the raw data, relying instead on standardized reports that the Office of Systems produced for them. The democratization of data was a very significant step in the advancement of data analytics in SSA. The OAO executives understood the data fields and how they were interrelated because of their knowledge of the disability process. In investigating issues, they also understood which data fields captured the most important and relevant information. Within a year of implementation of ACAT, OAO began to analyze the ACAT data.

To assist with the data analytics effort, OAO sought to hire data scientists. Unfortunately, human relations staff did not permit OAO to hire people with the necessary skill sets, arguing, essentially, that the OAO was a workload component and that data analytics belonged in a separate staff component. In effect, this approach separates data scientists from subject matter experts who understand and can readily explain what the data fields intend to capture, which data fields are most significance, and what the implications of the data anomalies represent. This solution also necessarily involves multiple chains of command with differing priorities.

Despite the pushback from the human relations staff, OAO was not dissuaded from establishing a data analysis team. One of the co-authors tapped data scientists around the agency and enlisted them to work with him in analyzing the data.[37] Having the authority to hire attorneys, OAO set about hiring a number of attorneys with data science degrees who later joined the data analytics efforts.[38]

OAO executives understood that compartmentalization of data science and operational tasks would not be a particularly nimble approach to effective data analysis. Data science and data analysis are a collection of highly-sophisticated mathematical processes. OAO executives understood that meaningful results are easiest to obtain when data scientists are paired with managers who understand the business process and identify problems that data could help solve. When the data scientists and managers collaborate effectively, operational managers can guide the data scientists toward the most relevant information, and the data scientists can transform the structured and unstructured information into readily digestible charts, graphs, and other visualizations that may reflect previously-uncovered operational realities. In turn, this enables the managers to use this information to make better decisions that improve the overall operation and delivery of services.

In OAO, the data analysis team performed six types of data analysis:

- **anomaly detection** – to identify observations that do not fit an expected pattern;
- **summarization** – to group or highlight information of particular interest;
- **regression analysis** – to model significant relationships otherwise hidden in the data;
- **classification** – to categorize where new observations most accurately belong;

---

37   The team was disbanded after the authors left their positions in what was then the Office of Disability Adjudication and Review.

38   One of the approved position descriptions in OAO is the position of attorney adviser.

- **dependency modeling** – to identify strong correlations between characteristics that can be used to predict outcomes; and
- **clustering analysis** – to subdivide data into smaller groups whose characteristics might help detect a structure or tendency that is not obvious while eliminating meaningless correlations and false suggestions.

OAO undertook extensive data quality cleaning before drawing conclusions from the data. Data fields are sometimes incomplete, reflect errors made when keying the data, or use fields that have been overwritten so that the values in the fields are different than they originally might have been. Data cleaning identifies and accounts for the errors and omissions in the data and also helps identify other opportunities for capturing valuable information not contained in the available data.

OAO used the data analysis in many ways, and their groundbreaking work is an example for other operational components to follow. In what might be described as an offshoot of queueing theory,[39] OAO staff worked closely with agency data scientists to develop clustering and multivariate probit analyses of its pending inventory to identify cases with similar characteristics in order to batch them for assignment, with the cases in each batch involving similar characteristics. It was believed this approach would work because the thought processes involved in adjudicating a case could be more readily and quickly repeated by assigning a batch of similar cases to be worked consecutively rather than working cases in random order. This modified case assignment effort resulted in nearly a 12 percent reduction in the time needed to process cases and a 7.5 percent improvement in work quality.

OAO staff also contracted data scientists to develop a naïve Bayes analysis of pending hearing level workloads that determined the probability of a finding of disability based solely on the characteristics of the claims. This analysis was not used to deterministically adjudicate the cases nor were the probabilities shared with adjudicators. The staff running the model advised managers of the cases with higher probabilities of allowance and the managers removed them from the pending workload queue and assigned them to adjudicators to be worked ahead of cases with lower probabilities of allowance under the notion that disabled claimants should receive their decisions as soon as possible.

Data analysis also helped identify that some claims warranting a finding of disability were being dismissed on procedural grounds. Another semi-naïve Bayes model was created to identify these cases so benefits could be rightfully awarded to this subset of claimants who, largely, were too impaired to meet the procedural demands of the hearing process. Obviously, those administrative procedures were designed to create and maintain an orderly appellate process and were not intended to deny benefits to those lawfully entitled to them.

The then-Acting Commissioner Colvin was so pleased with the results of OAO's data analytics efforts that in 2014 she established an academy[40] to teach other agency executives and personnel how to employ

---

39   Queueing theory is the theory of how service delivery efficiency can be improved by revising the order in which customers are served, often in a manner other than the traditional first-in, first-out order.

40   The Gerald Ray Academy, as Commissioner Colvin named it, is now run by the Analytics Center for Excellence.

the innovative methods used by OAO. In 2015, Acting Commissioner Colvin also created the Analytics Center for Excellence (ACE), a team of data science professionals tasked with helping infuse data-driven decision making across the SSA enterprise.[41]

In October 2017 and in continued furtherance of that effort, Acting Commissioner Nancy Berryhill moved the data analytics, anti-fraud efforts, quality assurance, and appellate review into one Deputy Commissioner led component, the Office of Analytics, Review and Oversight (OARO).[42] For the first time, the agency has consolidated the core components with the right skillsets to unlock the data and monitor the effectiveness of all of the operational components within SSA.

To take full advantage of the synergetic oversight opportunities this presents, OARO should alter the manner in which the ACE operates. SSA still has a tendency to assign data scientists to defined projects with specific timelines and deliverables. While there certainly is a need for such projects, the ACE data scientists should also be given wide latitude and time to engage in data discovery, reviewing and exploring current agency data, cleaning the data, and providing independent analyses of the information contained therein. Such an effort can be even more informative if the data scientists are co-located or otherwise work directly with operational personnel who can both describe what the data is intended to capture and identify operational problems that data analysis might help address. Data discovery by the ACE also may aid anti-fraud efforts. Data scientists and subject matter experts with an understanding of data science should work with the Office of Systems in the design of new computer systems to ensure that useful information is captured by those systems. The agency also should consider expanding data exchange agreements with other agencies, where possible, to improve fraud analysis and to further facilitate the deployment of deep learning algorithms in case analysis.

OARO also should team subject matter experts from OAO with the ACE to develop a contextual library of information to facilitate the use of deep learning algorithms in case work. Subject matter experts from OAO and data scientists from the ACE could assist quality assurance and appellate case review efforts by assisting with the design and development of better workload forecasting models, management information dashboards, and improved intelligent pathing in electronic case management systems and case analysis tools. Error types could be standardized, categorized, and classified to facilitate improved training and feedback and to help identify policy misalignments at all levels of adjudication. Additionally, the quality assurance work and training programs within OARO could be homogenized, adding greater focus on providing feedback to the initial and reconsideration levels at the individual level to improve performance there.

The authors recommend that OARO exploit the synergies created by the consolidation of the various components that make up OARO. The key focus should center on how the ACE is used. Co-locating subject matter experts with the ACE and allowing the ACE the time to undertake data discovery would enormously benefit OARO and ultimately the entire disability program.

---

41   The Executive Director's Broadcasts are published by the Social Security Administration, Office of Appellate Operations, and the information cited can be found in the Executive Director's Broadcasts article entitled ARPS Took OAO into New Era, published on January 6, 2017, at page 3. Copies available upon request at the Office of the Executive Director of OAO.

42   See the Social Security Administration Organizational Charts, at www.ssa.gov/org/oaro.htm. Last visited January 9, 2019.

The authors recommend that data scientists and subject matter experts work together to build better case analysis tools, forecasting models, and management information dashboards.

The authors also recommend that OARO adopt many of the quality assurance, training, and feedback techniques developed and used by OAO in its quality assurance efforts related to other aspects of the disability program, particularly at the initial and reconsideration levels.

III. The Expedience of Case Screening

Case Screening Efforts – the Appeals Council Process Improvement Initiative

Case screening can help reduce backlogs quickly if done correctly. OAO had a highly successful experience with case screening that began in March 2000 when it implemented the Appeals Council Process Improvement initiative (ACPI).

After productivity dropped precipitously following agency decisions in 1990 and again in 1992 to suspend the use of numeric-based performance rating expectations, including those related to productivity, OAO's backlog of cases began to grow. A few years later, the agency compounded the negative effects of eliminating productivity metrics by implementing a two-tiered, pass-fail performance rating system, which remained in place for many years. This rating system effectively removed any real performance measurement from job requirements.[43] The agency made little attempt to identify or acknowledge high performance. Instead, the agency adopted a generic non-measurable baseline for performance above which all employees passed regardless of their level of overperformance versus the baseline. Under this system, employees were encouraged to recommend themselves for awards. Many high performers, including the authors, found this offensive and declined to do so. Even more disturbing, almost no employees were rated unsuccessful under this standard, irrespective of how poor their performance might have been.

The pass-fail rating system, coupled with the suspension of measurable numeric-based productivity expectations, resulted in a widespread productivity decline that helped create a significant backlog of unworked cases. Processing times for OAO's appellate workload soared to an average of 555 days[44] by January 2000.[45] OAO was not the only operational component impacted by these decisions, although it may have been the only component that continued to capture data clearly showing the impact.

---

43   The decision to move to a pass-fail performance plan was taken despite the fact that the Corporate Leadership Council had earlier provided empirical evidence to the agency indicating that managers can have a substantial positive impact on unit performance if they ensure their employees understand performance standards and what is expected of them. *See* www.loc.gov/extranet/cld/workforce-performance/PerformanceDrivers.pdf. Last accessed December 21, 2018. *See also* https://marble-arch-online-courses.s3.amazonaws.com/CLC_Building the High Performance Workforce A Quantitative Analysis of the Effectiveness of Performance Management Strategies1.pdf. Last visited December 21, 2018.

44   For the entire fiscal year 2000, average processing time averaged 505 days; however, as measured month to month, it peaked at 555 days. Information about processing times at the Appeals Council are available at: https://catalog.data.gov/dataset/appeals-council-requests-average-processing-time/resource/b5c291b1-cdd2-4964-9556-e0125f216d32. Last visited April 5, 2019.

45   From 1989 to 1995, average processing times were under consistently below 240 days. For several years prior to 1992, branch managers were required to write a report describing the circumstances behind any case that remained pending for more than 120 days.

In the fall of 1999, one of the authors was selected to chair the ACPI initiative, an effort to address the mounting backlog. Barred from reintroducing productivity metrics into performance plans, the ACPI team used a different technique to drive higher performance: the implementation of differentiated case management.[46]

Notionally, the central idea behind ACPI was that there was no reason to make every claimant wait for an answer of an appeal until every case filed before his or her claim was processed.[47] Medical conditions, by their nature, can change or progress over time, and it is very common that new evidence is submitted at various times in the process. Requiring all claimants to wait for a very long time adds additional evidence to many of the claims, making appellate review more difficult. Not only is the record more voluminous, but the standard of review changes. When no new evidence has been received, the standard is primarily one of whether substantial evidence supports the hearing decision.[48] When new evidence is received, that standard remains, but the reviewing body, the Appeals Council, must also determine whether the evidence is new and material to the outcome of the claim. This additional adjudicative consideration adds time to adjudication. The solution was to triage the work by splitting it into two categories: easier cases with a more obvious likely resolution in one category and complex cases with a less obvious likely resolution in the other category.

The ACPI team tasked adjudicators and technical assistants with categorizing the cases. Analysts were invited to choose which type of work they preferred to process, but only 20 percent of the analysts at a time were allowed to process the obvious, easier cases, and then only if they maintained productivity of at least 4.0 cases per day.[49] The rest of the staff processed the more complex work in first-in, first-out order as they already were doing. The screening operation was quite successful in improving overall performance, reducing processing times, reducing the receipt of additional mail and evidence, and ultimately in driving down the backlog. The Appeals Council implemented ACPI beginning April 1, 2000, the midpoint of the fiscal year. Prior to implementation, during the first half of FY 2000, the Council processed approximately 43,000 cases. Following implementation of ACPI, during the second half of FY 2000, the Council processed more than 82,000 cases.[50]

---

46 Maureen Soloman of the School of Public Affairs at The American University in Washington, D.C., an early advocate of case-flow management improvement, made a presentation to OAO managers in the late 1990s. Her presentation influenced the ACPI team when it developed the differentiated case management approach to addressing the Appeals Council's request for review backlog.

47 There is no requirement that cases be processed in a first-in, first-out order. Other theories of docket management and the use of queueing theory can improve overall efficiency, but by employing these methods managers must be vigilant to ensure that no cases languish too long awaiting an agency decision or action. OAO shifted its focus from average processing time to ensuring that no cases remained pending longer than a specifically defined number of days and continually strived to reduce that defined number.

48 20 CFR 404.970 and 416.1470 provide the standards under which the Appeals Council will grant review.

49 The screening initiative portion of the ACPI plan was successful because it tapped into the intrinsic desire of people to excel at the work they perform. At the time, employees were under a pass/fail performance management system, so their level of productivity did not affect their performance rating. The productivity requirement attached to the screening initiative ensured that only highly self-motivated people would perform the screening and that they would maintain a high rate of performance.

50 The screening initiative of ACPI was quite successful in driving productivity higher by nearly 80 percent during the second half of Fiscal Year 2000. All but 5,122 of the cases processed in the second half of FY 2000 were processed by the existing OAO staff. Unfortunately, the productivity improvement largely went unreported by the agency or others who evaluated the effectiveness of the ACPI plan. The ACPI plan included several initiatives other than case screening that were either not measured or yielded only middling results. OAO leaders at the time did not fully execute the plan as it related to the processing of aged cases, so the processing time did not fall as much as the plan had projected. Although the backlog was significantly reduced, a significant portion of the drop in the backlog was a by-product of lower than expected receipts after Fiscal Year 2000. See "Appeals Council Process Improvement Action Plan," at https://oig.ssa.gov/sites/default/files/audit/full/pdf/A-12-02-12015.pdf. Last visited January 26, 2019.

Other Case Screening Efforts – The Senior Attorney Program: A Well-Intentioned Program that Lost Its Way

For several decades, the agency has struggled to secure a continuous supply of qualified ALJs from the Office of Personnel Management (OPM) to work down the case backlog. In light of this erratic supply line of ALJs, there was significant momentum to empower more experienced attorneys in the hearing offices to adjudicate and pay obvious allowance cases on the existing evidentiary record. These cases would be paid "on the record" without a hearing. The basic idea was that senior attorneys could triage and resolve these straightforward cases, many of which perhaps should have been awarded at earlier stages in the disability process, at the state agencies, in the first place. Meanwhile, the hearing operation would preserve precious ALJ resources for the more complex "gray area" cases.

This well-intentioned program functioned well at first. The program started slowly as the agency trained senior attorney adjudicators to approve cases without hearings. The Office of Quality Performance built models to help identify cases that might be awarded without a hearing, and the staff was tasked with screening cases identified by the models and issuing fully favorable decisions if such decisions were warranted. Although some quality assurance of this workload continued, over time oversight of this program waned. The limited controls on senior attorneys began to disappear, attorneys were encouraged to issue more decisions, and the numbers of favorable decisions the senior attorneys issued soared. As time went on, the global numbers jumped from several thousand cases paid on the record annually by senior attorneys to 54,186 cases in 2010.[51]

Why would a hearing operation begin paying such a large number of cases without a hearing? For one, there were significant pressures placed on the senior attorneys to approve cases more rapidly. Feedback to senior attorneys rewarded payment of claims over caution. Additionally, the lines between the agency and claimant representatives began to blur. At one point, the agency actively encouraged claimant representatives to provide lists of cases for "on-the-record" payment consideration. Many attorneys also provided computer disks that contained the wording for fully favorable decisions, ostensibly to save time for the hearing office in writing fully favorable decisions. By this time, quality control was virtually non-existent.

When the Appeals Council instituted own motion review of these favorable decisions, the Appeals Council learned that in many offices the models were not being used to identify the cases for possible award. The Appeals Council also found that a significant number of errors were being made.[52] The following year, quality assurance reviewers in the Office of Quality Performance, SSA's principle quality assurance component, reached the same conclusion.[53] The authors prohibited senior attorneys from soliciting candidate cases from claimant representatives. When the program came up for regulatory renewal, it was nearly jettisoned and barely survived in a significantly reduced state. Subsequently, OAO staff worked collaboratively with SSA economists to build a more accurate model, using a naïve Bayes analysis, to score the probability of allowance for each case in the hearing backlog. The new model was based on newer, richer data sources than were used in the original models.

---

51   See https://oig.ssa.gov/sites/default/files/audit/full/pdf/A-12-11018_0.pdf. Last visited January 9, 2019.
52   See https://oig.ssa.gov/audits-and-investigations/audit-reports/a-12-13-23002. Last visited January 21, 2019.
53   Ibid.

The newer data sources improved the accuracy of the models, but those data sources required quite a bit of data cleaning. Prioritizing accurate data capture at the initial and reconsideration levels would alleviate this problem and enhance the accuracy and value of the model. It could also lead to a more effective senior attorney program in addressing that portion of the hearing backlog that does not require a hearing. The enhanced data also has value for other data analytics projects related to the work done at the hearing level and Appeals Council.

The authors recommend that SSA maintain tight controls on all screening efforts with appropriate quality checks.

The authors recommend that SSA make accurate data capture by staff in operational components a priority. As necessary, SSA should include accurate data capture as a performance expectation in performance rating plans.

Other Case Screening Efforts – On-the-Record and Bench Decisions

In an effort to reduce the hearings backlog, OHO agreed to two initiatives that allowed ALJs to issue on-the-record decisions bench decisions. On-the-record decisions typically are issued before a hearing is held, obviating the delays that occur with the scheduling and holding of hearings. Bench decision are shortened-format decisions, sometimes issued from the bench at the time of the hearing. Both types of decisions can only be issued if the decisions are wholly favorable to the disability applicant.

Most of these decisions are issued by diligent ALJs who correctly analyze the records and identify when such decisions are warranted; however, a small subset of ALJs issue large numbers of on-the-record and bench decisions, and many of them have had quality issues. OAO's Division of Quality has done a good job identifying and resolving these problems through its focused review process, described in more detail *infra*. Nonetheless, like the senior attorney program, the on-the-record and bench decision initiatives could benefit from a structured data analysis of the characteristics of the underlying claims and statistical modeling to help identify the cases most likely to be correctly decided using these decisional formats.

The authors recommend the agency conduct ongoing data analyses of the characteristics of pending workloads to identify the cases that can be processed under the senior attorney, on-the-record, and bench decision initiatives.

The authors further recommend that the quality of these decisions be carefully monitored on a regular and continuing basis and that the opportunity to produce these types of decisions be revoked until training is completed whenever significant quality concerns are identified.

IV. Toward Improved Quality Assurance

In terms of quality assurance of the hearing level work, the agency seems to have a preference to use national sampling that provides little insight into the quality of work at the individual level. In the past, the sample size of the quality samples was not sufficient to review even one case from all of the adjudicators on an annual basis. An overall quality metric was statistically derived based on the errors found in the sample, and the agency reported this percentage as an overall indicator of quality at the hearing level on an annual basis. Congressional oversight committees and Office of Management and Budget staff seemed to accept the high levels of quality that were consistently reported.

Of interest to the authors, however, was the fact that the needle rarely moved in terms of the accuracy reported, and the reported accuracy rates seemed high based on their personal experience. The authors realized that these consistently high accuracy rates implied that leadership changes, policy changes, training initiatives, technology modernization, and staff turnover must have little effect on overall quality. Doubtful about these potential ramifications, the authors implemented other approaches to quality assurance.[54]

The authors believed that the point of quality assurance is to identify where errors are being made so that corrective action can be taken to address the root causes of the errors. Reporting the results is simply a by-product of that exercise. Of particular immediate interest was the newly available structured information about the specific reasons why cases were being remanded from both the Appeals Council and the Federal courts. The data was obtained through use of the ACAT. Analysis of the data clearly identified which cases were being remanded and reworked and why. It also clearly identified whose adjudicative work was the subject of the remands and rework. OAO paired this analysis with other data identifying outlier behaviors.

The authors were also perplexed as to how any ALJ could pay benefits to nearly 100 percent of the claimants appearing before him or her when the cases were supposed to be randomly assigned. Similarly, top national leadership was confounded as to how any judge could pay less than 10 percent of his or her cases. To be sure, there were close to a dozen of these low-paying ALJs nationwide too. But at least these low-paying ALJs would receive some oversight by the Appeals Council and/or Federal court in many instances. In contrast, the generosity of the high allowance ALJs was virtually unreviewable based on the program design at that time. Successful grants of benefits were reviewable only under the own motion authority of the Appeals Council, which was rarely used.

To investigate the patterns that were emerging, one of the authors persuaded then-Commissioner Michael Astrue to allow OAO to reposition some of its existing staff for this purpose through *sua sponte* reviews of unappealed claims.[55] In October 2010, OAO stood up the Division of Quality, carved from existing staff

---

54 According to an Executive Director's Broadcast, OAO has reported that they are resuming national sampling. The Executive Director's Broadcasts are published by the Social Security Administration, Office of Appellate Operations, and the information cited can be found in the Executive Director's Broadcasts article entitled See DQ, OQR Work Together to Expand Quality Review, published on October 5, 2018, at page 3, Copies available from the OAO Executive Director's Office upon request.

55 The Executive Director's Broadcasts are published by the Social Security Administration, Office of Appellate Operations, and the information cited can be found in the Executive Director's Broadcasts article entitled OAO Makes History with Launch of Quality Review Branches, published on September 17, 2010, at page 3. Copies available from the OAO Executive Director's Office upon request.

who had become more productive as a result of the earlier implementation of productivity standards and goals as well as refined training on how to do the work efficiently.[56] The authors tasked the Division of Quality with sampling unappealed favorable decisions and dismissals and capturing structured data about the findings of those reviews. In 2011, they also put in place a system of "focused reviews" that reviewed cases involving specific issues or work produced by or involving certain individuals related to Social Security programs or on the work of any person who processes claims or provides evidence to the agency. Data scientists worked with OAO staff to develop a methodology for prioritizing focused reviews. Of particular interest were ALJs exhibiting outlier behaviors, usually two or more standard deviations from the mean, in some aspect of their work.

All focused reviews were conducted by OAO analysts experienced in case review. Once the review was completed, the reviewers presented their findings to headquarters executives who arranged with regional and local management ALJs to provide feedback directly to the individual ALJs.

This program was controversial at first but soon became accepted practice, albeit begrudgingly by some in the ALJ corps. The reviews were intended to generate useful feedback for ALJs who previously had rarely received any. The limitation of focused reviews was apparent from its title though – the reviews honed in on specific issues and decisions of only a small number of ALJs or issues each year.

A few obviously irregular patterns emerged from the data, and focused reviews and were referred to the Office of the Inspector General for further investigation. An obvious pattern of irregularity was identified by OAO's Division of Quality and OHO data scientists related to one of the ALJs in the Huntington, West Virginia hearing office, although it was not the genesis of the investigation by the Department of Justice that led to several indictments and jail sentences.[57] Generally, however, careful analysis of the data and information from quality assurance and focused reviews reflected that most of the outlier behaviors the reviews uncovered did not appear to have a fraud component. Instead, it was clear to OAO leadership that many remands resulted from divergence from the policy-compliant pathing in the analysis of the disability claims. They surmised that the divergence likely resulted from the use of incomplete heuristic models[58] individually developed by some adjudicators. Drawing on the work of Princeton Professor Daniel Kahneman,[59] the Council sought ways to improve feedback to adjudicators to ensure that the heuristics they individually developed and used in case adjudication also complied with agency policies. The theory was that nearly all ALJs strive to do the right thing and would appreciate assistance and feedback to help them do so.

---

56   The Division of Quality began with about 60 analysts, five clerical support staff, five branch chiefs, and ten adjudicators. The Executive Director's Broadcasts are published by the Social Security Administration, Office of Appellate Operations, and the information cited can be found in the Executive Director's Broadcasts article entitled OAO Makes History with Launch of Quality Review Branches, published on September 17, 2010, at page 3. Copies available from the OAO Executive Director's Office upon request.

57   See the Appendix to this paper for a discussion of the Huntington matter.

58   Heuristics can be considered as rules of thumb people develop and use to process complex information.

59   Dr. Kahneman, the 2002 Nobel Memorial Economic Prize in Economic Sciences awardee, has suggested that heuristics can be improved if people are immersed in a business process, the business rules don't change frequently, and the individuals are provided with feedback about their performance. See Daniel Kahneman, Thinking Fast and Slow, Farrar, Straus and Giroux, New York, New York, 2011.

Historically, hearing office management would alert each ALJ about their production each month. Noticeably absent in this feedback was any indication about the quality of their decisions. The feedback from the focused reviews proved to be a very effective method for providing direct information to ALJs about the quality of their work. Due to the limited number of these reviews, however, it was not an effective method for providing feedback to the vast majority of ALJs about their work quality. Clearly, there was a glaring need for some other type of quality metric.

At the time, the Appeals Council was capturing a national grant review rate, which reflected the percentage of appealed cases in which the Appeals Council granted review to take corrective action or return cases via remand order to the hearing offices for further consideration. Similarly, the Appeals Council captured its rate of own motion review of hearing decisions at a regional and national level.

Applying the conceptual framework of these quality metrics, the authors established the imperfect but straightforward metric known as the agreement rate.[60] In short, this was the percentage of cases the Appeals Council reviewed in which it did not remand, or reverse the outcome of, each particular ALJ's reviewed decisions,[61] for reasons within the control of the ALJ.[62] When this effort began in 2012, the ALJ agreement rate was in the mid-70s. In other words, the Appeals Council did not disagree with more than 70 percent of the ALJ decisions that it reviewed nationwide. This figure would rise significantly in subsequent years. Own motion and grant review rates measuring the percentages of cases in which the Appeals Council reviewed and took corrective action conversely fell in subsequent years.

With a backlogged system focused on production historically, though, there was significant distrust among the ALJ corps when the hearing operation moved to a more nuanced and balanced working model that valued quality too. This was a major directional shift inasmuch as prior shifts towards quality had been short-lived and quickly abandoned by upper management.

The next challenge was to make quality, productivity, and a host of other data and information available to the management team and line ALJs in real time. The authors, with great support from their information technology colleagues, developed the previously described web portal known as "How MI Doing."

This portal provides a host of production, quality, and hearing office related information to every ALJ and management official at their respective desktop. It is gated, based on the need-to-know. ALJs have access to data and metric measurements related to the cases they have been assigned or have processed, as well as comparative aggregated data about how local peers, regional peers, and national peers were performing. Many ALJs were shocked at their first glance at the new quality metrics, updated daily and readily available from their desktop computers. But nobody could dispute that this infusion of data moved the bar on quality decisions in the hearing operation. By 2015, the agreement rate was approaching 90 percent, up from 70 percent five years earlier.

---

60   The term "agreement rate" is a bit of a misnomer. The Appeals Council does not decide whether it agrees or disagrees with the ultimate decision on the issue of disability but instead only grants, denies, or dismisses review of an applicant's appeal of a hearing decision after deciding whether it has a legal basis for granting review.

61   One flaw with this metric, much like the reported national random sample accuracy rate, is that the Appeals Council does not review significant numbers of cases decided by each ALJ each year.

62   Certain categories of remanded cases are excluded from the calculation of the "agreement rate." The cases are excluded because sometimes cases are remanded based on information not available to the ALJ at the time the hearing was issued, such as the post-hearing decision receipt of new and material evidence.

The authors recommend regular review of office and individual-level data for unusual patterns and practices across all hearing offices nationwide.

The authors recommend the creation of a nationwide health care provider database containing structured data about and unique identifiers for every medical and mental health care provider who submits evidence and/or reports in support of disability applications. Ideally, this database would include direct links to every document submitted to the agency from any health care provider.

The authors recommend that SSA executives and managers emphasize operational productivity and quality metrics equally.

The authors recommend reconstituting a nationwide hearing office litigation team in the Office of the General Counsel. This group should be staffed sufficiently with experienced employment litigators from across the country.

The authors recommend that line hearing office managers be required to review local performance data and subsequently report progress and any anomalies to senior management.

V.   A Look into the Future – Boolean Policy and Computational Law

Toward Improved Clarity in Policy Guidance

Adjudication in general can be described in terms of propositional logic if the steps in the adjudication are dissected sufficiently.[63] After nearly 60 years of mass adjudication, there are very few actually novel issues that arise in Social Security disability claims. Over the years, adjudicators have used inquiry and analysis to break initially broad issues into a series of micro-issues upon which they make their findings. Most of these micro-decisions have been considered countless times by many adjudicators, and patterns have emerged in remanded cases in the areas in which adjudicators act inconsistently with one another. Since policies are intended to reach consistent adjudication in cases with similar evidence and issues, it follows that policy guidance could be crafted to guide adjudicators toward or through the steps of the key propositional logic micro-decisions.

It was this approach that enabled one of the authors and his colleagues at the Appeals Council to develop a refined decision tree of the steps necessary for proper adjudication that leads to the approximately 2,000 types of decisions that can be issued in disability claims. The decision tree was crafted in a way that might be described as Boolean – the policies were mapped using propositional logic in a manner similar

---

63   Law schools inherently recognize the underlying importance of propositional logic as they consider LSAT results as one of the key factors in the admission process. A significant portion of the LSAT, the analytical reasoning or logic games section, can be easily mastered with the use of Venn diagrams, which effectively are based on propositional logic.

to the logic processes used in Boolean algebra.[64] The more precisely regulations, rulings, and other sub-regulatory policies, procedures, and guidance are drafted in a way that pushes adjudicators toward these Boolean-type outcomes, the easier it is to model the policies and develop machine learning and deep learning algorithms for use in decisional assistance tools that can be used in helping determine whether the policies are followed. This approach also improves the likelihood that adjudicators consistently will follow the policies correctly and thus can be expected to reduce remands.

At the heart of the Social Security disability evaluation process is a structured approach known as the sequential evaluation process,[65] which already guides adjudicators through a series of somewhat Boolean inquiries.[66] At the first step, there are two deterministic and two non-deterministic outcomes: the claimant is engaging in substantial gainful activity (SGA) as an employee or as a self-employed individual and therefore is not disabled, or the claimant is not engaging in SGA in either of these capacities and the sequential evaluation process proceeds to the next step.[67] At the second step, a deterministic finding of not disabled obtains if the claimant's conditions, considered separately and in combination, are not severe.[68]

Adjudicators proceed to the third step in the sequential process[69] if one or more severe impairments is identified. Hundreds of Boolean, outcome-determinative conclusions are possible at the third step of the process, where the medical signs and/or laboratory findings match or are found to be equivalent in severity to pre-defined medical conditions that are deemed to be *per se* disabling without consideration of the educational and vocational characteristics of the claimant.[70]

If the claimant's conditions do not meet or medically equal any of the listed impairments, the sequential process directs adjudicators to establish the claimant's residual functional capacity (RFC), i.e., the work-related functionality the claimant retains after consideration of the effects of the claimant's medical conditions. The regulations specify that RFC is to be cast in terms of the claimant's maximum capacity for performing work-related activities.[71] This information is necessary because beyond the listed impairment step in the sequential evaluation process adjudicators must consider vocational factors in deciding disability.

SSA provides clear guidance related to the evaluation of the physically-exertional aspects of RFC. SSA generally categorizes exertional limitations in increments, with lifting and carrying weights considered in ten-pound increments; sitting, standing, and walking capacities measured in hours out of an eight-hour day; and so on. Additionally, those increments have been sorted into broader buckets defined regulatorily as very heavy, heavy, medium, light, sedentary, and less than sedentary work. These categorizations make it relatively easy to develop vocational information correlating the most common exertional demands of work with the exertional functional capacities of disability applicants.

---

64   Policy formulated under this approach would held guide adjudicators toward decision points that generally could be decided in the form of choices, such as yes/no, true/false, if/then, and/or and not, or similar constructions, in much the way choices are forced in Boolean algebra.
65   Specific variations apply to certain types of claims, such as claims of children, widows, widowers, and blind applicants.
66   The general disability evaluation process is described in 20 CFR 404.1520ff, and 20 CFR 416.920ff.
67   See generally 20 CFR 404.1571-1576, and 20 CFR 416.971-976.
68   See 20 CFR 404.1521 and 416.921.
69   See 20 CFR 404.1525-1526, and 20 CFR 416.925-926.
70   The listed impairments can be found in Part 404, Subpart P, Appendix 1 to SSA Regulations No. 4.
71   See 20 CFR 404.1545-1546, and 20 CFR 416.945-946.

The language relating to mental impairment severity often couches severity in terms of extreme, marked, or moderate limitations of functioning. Many of these severity terms are ill-defined and are difficult to translate into the precise maximum RFC a claimant might have.

SSA's regulations and a number of Social Security Rulings do offer some guidance about how non-exertional limitations are to be considered in evaluating RFC. Mental impairments,[72] for example, generally are evaluated in terms of the ability to understand, remember, or carry out simple instructions; use judgment; respond appropriately to supervision, co-workers, and usual work situations; and cope with changes in a routine work setting.[73] In recent years, adjudicators also have seen an increase in allegations from claimants that their mental impairments cause an inability to travel to work or inability to attend work on a regular basis.

The agency guidance regarding the evaluation of the non-exertional components of RFC is rather open-ended in terms of how degrees of limitations affect degrees of remaining functioning. Combinations of possible non-exertional limitations and exertional limitations lead to billions of potential RFC permutations. This makes it difficult for adjudicators to precisely formulate how the varying degrees of limitations affect the ability of the claimant to function in a job setting, and this can lead to haphazard and inconsistent decision-making.

There are three obvious and logical points in the sequential evaluation process where some clarity could be brought to bear on the evaluation of non-exertional impairments – at the listing step, at the point of RFC formulation, or at the point where vocational information is introduced into the evaluation. SSA might consider attempting to address most non-exertional limitations in a binary fashion at step 2 of the sequential evaluation,[74] but such an approach likely would be controversial. SSA could develop and codify broad categorizations of non-exertional limitations in the same manner as they have done with exertional limitations, although it is difficult to imagine how the agency would do so.[75] Most of the agency's current efforts are focused on the third path – the development of improved and more current vocational information.

Without going into a detailed discussion of the complete sequential evaluation and disability evaluation process, once the agency proceeds beyond the listing step the agency must compare a claimant's RFC

---

72 SSA also provides regulatory and sub-regulatory guidance concerning the evaluation of other non-exertional limitations. See generally 20 CFR 404.1569(a) and 416.969(a) and related Social Security Rulings, found at https://www.ssa.gov/OP_Home/rulings/rulfind1.html. Last visited April 30, 2019.

73 See 20 CFR 404.1545 and 416.945.

74 Such a notion might imply, for example, that claimants with certain mental impairments found not to be of listings severity or equivalent also would be found capable of traveling to and attending work on a regular basis and capable of meeting the mental demands of unskilled work, including the ability to concentrate sufficiently to attend to tasks, follow simple instructions, and interact appropriately with coworkers, supervisors, and the public. Alternatively, a scaled score for severity, such as the scaled score used in evaluating intellectual disability, might be workable, but ultimately there would still be a point on the scale indicative of disability and a point on the scale indicative of no disability. Creating a gray area in between those points could lead to inconsistent adjudication, whereas a binary approach would not. Similarly, claimants with pulmonary, vision, hearing, skin, and other conditions that create certain specified non-exertional limitations who are not found disabled at step three might, for example, be found to have the non-exertional capacity to perform unskilled work not involving moving or dangerous machinery or exposure to significant environmental irritants.

75 Considering mental functioning by hours in the day hardly seems workable. Is it reasonable to conclude a person can perform full-time substantial gainful activity in a non-sheltered work environment if, for example, they can only interact appropriately with coworkers and supervisors up to five hours per work day, or only concentrate on job tasks four hours per work day?

with the functional demands of past relevant work and other work existing in significant numbers in the local or national economy.[76] When considering all of the possible limitations that can be caused by the myriad of medical conditions that a claimant might have, it becomes obvious that there are billions of possible RFCs. To complete the sequential evaluation process, the agency must have access to information sufficient to compare these billions of possible RFCs with the millions of different types of jobs in the national economy. Granular information addressing the effects of all of the possible limitations on the work demands of all jobs is simply not available. These problems are further compounded by the fact that the nature of work is undergoing significant transformation. There is a strong likelihood that many job functions, and some jobs, may be displaced in the near future because of the rapid development and expansion of the use of artificial intelligence in the business world.[77] Traditional notions of work are also being upended by the rapid growth of the gig economy, in which nearly a quarter of the workforce participate in part-time work for remuneration.[78]

One of the primary sources of vocational information used by the agency is the Dictionary of Occupational Titles, a vocational source produced by the Bureau of Labor Statistics (BLS) that was last published in 1977, with the last minor update produced in 1991.[79] The agency will need access to significant micro-data about the current and frequently changing job market to ensure that its disability decisions are amply supported with robust job information. To that end, the agency has contracted with BLS to conduct new surveys and update this critical information. Information about the exertional demands of work is relatively straightforward and easy to obtain, but given the myriad possibilities of how non-exertional limitations can affect job functioning, the task of obtaining data about those limitations is much more onerous. The project has been ongoing for many years and still is not near completion. The task is Herculean and litigation risks are likely intensifying. It is imperative that the agency quickly develop a viable solution to the lack of adequate current vocational information to support its decisional conclusions.

While acquiring updated information about job demands through the BLS survey will likely help,

---

76 Once the adjudicator establishes an RFC for a claimant, the adjudicator must consider whether a claimant can perform past relevant work (PRW). See 20 CFR 404.1545, 404.1560, 416.945 and 416.960. If the claimant meets the burden of establishing an inability to perform past relevant work, the agency has an incumbent duty to identify jobs existing in significant numbers in the local or national economy that the claimant could perform. Adjudicators do this by comparing the claimant's RFC with the functional requirements of jobs in the local and national economy to ascertain whether the claimant retains the capacity to perform such jobs that exist in significant numbers. At the heart of this step in the sequential evaluation process is a series of vocational Grid Rules and vocational profiles that filter the billions of possible residual functional capacity conclusions, considered in the context of the claimant's age, education, and transferability of skills from past relevant work experience, into the binary conclusions of disabled or not disabled. The ultimate conclusion of disability at this step is then determined by the age and educational attainment of the claimant, with illiteracy and the inability to communicate in English as additional factors for some individuals who are not of advanced age, although SSA has recently proposed revising the rules related to the inability to speak English. See https://www.federalregister.gov/documents/2019/02/01/2019-00250/removing-inability-to-communicate-in-english-as-an-education-category. Last visited April 29, 2019. For a more complete discussion of these factors, see 20 CFR 404.1560-1569 and 416.960-969. At the hearing level, ALJs often solicit testimony from vocational experts to support the conclusions reached, particularly when the Grid Rules cannot be directly applied or when transferability of skills is at issue.
77 https://www.brookings.edu/research/automation-and-artificial-intelligence-how-machines-affect-people-and-places/. Last visited April 16, 2019.
78 https://www.forbes.com/sites/tjmccue/2018/08/31/57-million-u-s-workers-are-part-of-the-gig-economy/. Last visited April 16, 2019.
79 The agency captures a wealth of data from disability applicants about work they performed in the past. In 2013, one of the authors proposed cross-matching and combining existing job data from the current DOT, the U.S. Census Bureau, and other sources with data about jobs captured by the agency from disability applicants to refresh and improve the overall quality of data then being used in disability decisions. He also proposed expanding and improving the data-capture forms the agency currently uses to improve the quality and usefulness of the data it captures. To date, the agency has not shown much interest in pursuing these ideas.

redesign of the policies related to the consideration of non-exertional limitations would also be beneficial. Irrespective of whether SSA creates broad categorizations for non-exertional limitations, addresses these limitations using a binary choice method at step 2 of the sequential process, obtains better data about the effects of non-exertional limitations of work functioning, or crafts some other solution, it seems obvious that better-defined policy for addressing non-exertional limitations would improve the consistency and quality of agency adjudication. Improvements in this area of policy would also help ameliorate existing vocational policy issues and aid any effort to develop deep learning and artificial intelligence algorithms in decisional support tools.

The authors are persuaded that the best path to a solution begins with data analysis.[80] OAO used data analysis and data visualizations to uncover policy issues and push policy solutions related to the evaluation of subjective complaints, medical opinion evidence, and submission of evidence, issues that previously had resulted in large numbers of remands by both the Appeals Council and the Federal courts. SSA could undertake extensive data analysis of existing data related to disability claims adjudicated by ALJs that involved impairments that produce non-exertional limitations to identify the circumstances in which the non-exertional impairments were a determining factor in whether the claims were awarded or denied.[81] OAO's Division of Quality could conduct large samples of these cases to identify and cull specific information about the differences between those correctly allowed and those correctly denied. Such an effort would greatly aid any analysis of non-exertional policy and support potential changes that would benefit agency adjudication.

The authors recommend that SSA undertake careful data analysis and case review to determine how best to address policies related to the evaluation of impairments that cause significant non-exertional limitations.

Computational Law and the Future of Decisional Assistance

Social Security disability decisions are written by hundreds of different paralegals and attorneys and reviewed by hundreds of ALJs. The language used by so many different people varies significantly even when the issues are quite similar. Despite the differences, however, much of the terminology in decisions is used as terms of art with specific meanings in the context of agency regulations. The decisions follow a known format, and staff at the Appeals Council previously had mapped the policy-compliant pathing of that decisional format when ACAT was constructed. The effort reduced dimensionality problems related to the application of disability policies, and the data from ACAT and the knowledge of the policy-

---

80 A concern about using data analysis of disability claims to underpin agency disability policy is that the agency might risk applying general information to individual claims. The authors understand that mass adjudication, after all, is actually a long series of individual adjudications and that each claimant is entitled to individual adjudication. But patterns emerge from the sheer volume of cases adjudicated, and individual cases are distinguished based on individual differences in the fact patterns. The agency already recognizes that many patterns exist and have used these patterns to describe the exertional requirements of jobs. They also have used these patterns to not only direct conclusions of disability, both at the listing step and in the direct application of Grid Rules, but also to direct conclusions denying disability claims under the direct application of Grid Rules. Thus, the authors believe that patterns can be discerned and used by the agency to structure the evaluation of non-exertional limitations, provided that claimants retain the opportunity to distinguish the facts of their individual cases from the patterns to forestall application of those patterns to cases in which the facts do not square with the patterns.

81 Another possible approach would be for the Appeals Council to issue a series of precedential decisions, an idea floated by many others in the past. The authors believe that legislative changes or rulemaking would enable a fuller and richer discussion of the issues by the public and elected officials.

compliant pathing provided an opportunity for enterprising OAO attorneys to build an analytical tool capable of evaluating decisional language in the context of regulatory sufficiency. Using optical character recognition technology, information technology specialists converted decisional documents, maintained in TIFF format, to a text-mineable format. OAO subject matter experts successfully converted the context rich language of the hearing decisions into the logic-based rules of natural language processing in an analytical tool they developed known as INSIGHT.[82]

Natural language processing (NLP), also known as computational linguistics, uses mathematical algorithms to evaluate natural language texts through pattern recognition and computer heuristics. To facilitate the use of NLP, OAO staff manually reviewed thousands of hearing decisions in a word sense disambiguation exercise designed to create a training set of information for use by the algorithms underpinning INSIGHT. The staff manually curated the data and information behind the algorithms and validated the results. The analysts identified language constructs that address issues that frequently result in remand and those that did not.

The tool they developed extracts language from hearing decisions and analyzes whether the information imparted by that language conforms with existing agency policy. It also checks decisional language for internal consistency. INSIGHT does this by using machine learning algorithms to probabilistically identify commonalities in existing data and make predictions about new data. Word sense disambiguation enabled INSIGHT to use a string search algorithm known as regular expression to identify and extract relevant terminology and decisional language. INSIGHT applies lexical semantics algorithms to determine the meaning of words within the context of their use, fuzzy string matching to identify spelling variants and typographical errors, and co-reference resolution to find other words, phrasing, and notations with essentially the same meanings. These types of algorithms enable INSIGHT to identify multiple constructs of the same basic ideas. INSIGHT also uses named entity recognition classification to identify proper names and make deductions about the type of each name, i.e., whether the name is that of a person, organization, or place.

A hearings and appeals analyst can run INSIGHT from a desktop application. INSIGHT can scan an entire hearing decision in a few seconds and correctly identify more than one-third of the issues that generate Appeals Council remands with a high degree of accuracy. After scanning a decision, INSIGHT produces a short report of findings highlighting potential problems in the decision reviewed. It includes hyperlinks to the specific problematic language in the hearing decision and to related policy guidance. The analyst reviews the INSIGHT output and then recommends to an AAJ what, if any, action the Appeals Council should take. The tool also includes an email generating feedback mechanism if an analyst identifies an error, omission, or other problem in the INSIGHT tool or the output it supplied. Recently, OHO has begun testing the use of INSIGHT in hearing offices, where decision writers can run the INSIGHT tool against decisions they draft to highlight issues they can resolve prior to issuance of the decision by an ALJ.

INSIGHT does not assist adjudicators in navigating the medical records or in making judgments related to the evidence. The opportunity for development of a tool that could perform such tasks now exists,

---

82   INSIGHT was the brain-child of attorney adviser Kurt Glaze, who worked closely with Appeals Officer Robyn Konkel in particular, as well as many others in OAO to develop the tool.

although progress will be slow because the vast majority of the hundreds of millions of pages of medical and vocational evidence potentially available in these disability claims are stored as electronically as TIFF images, while 4 percent of pending disability claim files remain as paper records. In the short term, some of the TIFF images can be converted to text-mineable format. The future likely will involve the collection and storage of medical information in text-mineable form. Indeed, the agency already has begun collecting and storing disability claim files in structured data format in a small but growing percentage of claims (currently less than 5 percent of all pending claims).

Currently, the agency is pursuing artificial intelligence efforts designed to highlight for case reviewers, in a non-deterministic way, language in medical reports that correlates well with language in the medical listings to presumptively find disability. While perhaps of some utility, the authors believe such an approach is not entirely consistent with the nature of adjudication, which requires the evaluation and weighing of often inconsistent information. Highlighting only some of the information in the record could lead some time-burdened adjudicators to inappropriately rely on the information the algorithms uncover. If not carefully monitored by expert adjudicators, this could result in an uptick in improper payments.

The use of artificial intelligence in reading and interpreting medical records is promising, and may prove to be a great boon to improving the speed, consistency, and quality of decision making. There are three core issues that must be considered before building artificial intelligence platforms intended to be used in a deterministic manner to find disability. The first of these issues relates to the act of adjudication itself, which requires a weighing of the evidence before a final decision is made. Nearly all records have information that could lead to conflicting outcomes on the issue of disability, and the Social Security Act requires that the evidence be weighed and the conflicts carefully considered. For example, deterministic algorithms would need the capacity to not only find language correlating with language in the listings, but also would need to the capacity to identify language that is not consistent with the listings. The algorithms also would need a mechanism for identifying and addressing the conflicts in the record.

Transparency of the decision-making process is the second core issue related to using algorithms in a deterministic manner. Claimants, their representatives, and the federal courts expect agency decisions to explain how core findings and the ultimate decision have been reached. While deterministic algorithms can be tied to language describing the findings and conclusions of hearing decisions, questions may still arise as to whether the algorithms have correctly considered the evidence in reaching those findings and conclusions. To an outsider, the algorithms can appear to be black box technology, and it is highly likely that the construction of the algorithms and how they operate eventually will be subjected to judicial review with the very real possibility that they could be held to be arbitrary and capricious.

If the agency is able to overcome issues raised related to the transparency in decision-making, the agency likely will face a third core issue: the possible characterization of the use of deterministic algorithms as administrative overreach that subverts the rulemaking process. If used in a deterministic fashion, the algorithms, in effect, would be applying rules to make findings and an ultimate conclusion. The issue would relate to the extent to which the public has a right to review, or has had an opportunity to review and comment upon, the rules used by the algorithms.

To address these legal issues, the Appeals Council, with help from the ACE, began investigating the construction of a Bayesian belief network[83] that would help SSA structure and categorize the data and information necessary for deep learning and artificial intelligence algorithms. Such algorithms would be capable of reading and extracting information from the medical records associated with disability claims, but the Bayesian belief network would provide a level of transparency to show how the algorithms interpret and understand the records within the context of the Social Security Act, its implementing regulations, and agency policies. Agency publication of the data and descriptions of how the data is used by the algorithms might be sufficient for the agency to withstand the types of legal challenges mentioned herein. The agency chose not to support the Appeals Council's effort to construct a Bayesian belief network, and the authors understand that effort has ended. Nonetheless, the authors believe that a legally viable approach exists for building artificial intelligence powered algorithms for use in a deterministic manner in the disability program. The authors believe such algorithms would be particularly useful at the initial and reconsideration steps in the disability program. Any algorithms should have transparent and accessible supporting data tables with documentation establishing the clear linkages of the algorithmic outcomes to Boolean-constructed policies that have been carefully developed through the rulemaking process. The authors believe that machine learning, NLP and artificial intelligence hold great promise for assisting the agency in addressing its workloads.

*The authors recommend that the agency carefully consider the legal issues described herein before investing too heavily in its current or future artificial intelligence-related effort*s.

---

83   What the Appeals Council characterized as a Bayesian belief network essentially was a graphical representation of the array of information extracted for use by the machine learning/NLP program, the conditional dependencies of the arrayed information, and the conditional probabilities of various outcomes based on the algorithms applied to that information.

Appendix: Lessons from the Huntington Hearing Office Fraud Case

The story of the Huntington, West Virginia Hearing Office (Huntington) has been told from any number of vantage points. National newspapers reported that it was the largest disability fraud case in the history of the disability program.[84] Local news outlets opted for the personal angle by focusing on the scheme's mastermind.[85] News magazines highlighted acts of greed and arrogance committed by various members of the Bar. [86] Even the FBI got into the act, urging the public to help find the flamboyant lawyer who fled to Honduras via Mexico just hours after admitting guilt to prosecutors.[87] Stripped down to its basics, a corrupt ALJ colluded with a corrupt legal representative in a small town with plenty of help from a few members of the local medical community who submitted thousands of falsified medical documents.[88]

The longevity and scope of this fraud case make this case worthy of examination. According to admissions as part of attorney Eric Christopher Conn's plea, this scheme lasted for over a decade at an estimated cost to taxpayers of over $550 million.[89]

In light of the magnitude of the scheme, massive dollar loss to taxpayers, and the reputational hit to SSA's programs, the participants received lengthy sentences for these white-collar offenses. For example, the mastermind Conn received a 12-year sentence and an additional 15 years for fleeing prosecution.[90] ALJ David Daugherty,[91] at age 81, was sentenced to four years in prison.[92] A clinical psychologist, Alfred Bradley Atkins, who submitted fictitious evidence to SSA was found guilty after a six-day trial and was sentenced to 25 years in prison.[93]

While the criminal cases received extensive coverage, little press attention was devoted to how this massive disability fraud was perpetuated in this sleepy little town. What were the programmatic conditions that set the stage? Where was regional management and local oversight for this far-flung outpost? Did the

---

84   A Department of Justice (DOJ) press release from September 7, 2018, states that this case was "the largest fraud scheme in the history of the Social Security program." See "Fugitive Lawyer in Largest Social Security Fraud Scheme Sentenced to 15 Years in Prison for His Escape and Related Crimes," DOJ Press Release 18-1159, September 7, 2018, available at https://www.justice.gov/opa/pr/fugitive-lawyer-involved-largest-social-security-fraud-scheme-sentenced-15-years-prison-his. Last visited April 18, 2019.

85   See, for example, Bill Estep, "A really odd person": Fugitive lawyer Eric C. Conn's bizarre life story." Published June 23, 2017 by the Lexington Herald Leader. Available at www.kentucky.com/news/state/article157933819.html. Last visited January 29, 2019.

86   See "Disability, USA," first aired by CBS News on October 6, 2013, available at https://www.cbsnews.com/news/disability-usa/. Last visited April 18, 2019.

87   DOJ Press Release 18-1159, September 7, 2018, op. cit. Last visited January 27, 2019.

88   Ibid. See also United States v. Conn, Daugherty, and Atkins, Indictment No. 5:16-Cl-22-DCR (April 1, 2016) (detailing scheme to provide false medical evidence to SSA).

89   DOJ Press Release 18-1159, September 7, 2018, op. cit. Last visited January 27, 2019.

90   Ibid.

91   In admissions made as part of his guilty plea, Daugherty admitted to receiving more than $600,000 in cash from Conn for approving over 3,000 of Conn's client's cases. See "Former Social Security Administrative Law Judge Sentenced to Four Years in Prison for Role in $550 Million Social Security Fraud Scheme," DOJ Press Release 17-941, August 25, 2017, available at https://www.justice.gov/opa/pr/former-social-security-administrative-law-judge-sentenced-four-years-prison-role-550-million. Last visited April 18,2019.

92   DOJ Press Release 18-1159, September 7, 2018, op. cit. Last visited January 27, 2019.

93   "Former Clinical Psychologist Sentenced to 25 Years in Prison for Role in $550 Million Social Security Fraud Scheme," DOJ Press Release 17-1046, September 22, 2017, available at https://www.justice.gov/opa/pr/former-clinical-psychologist-sentenced-25-years-prison-role-550-million-social-security-fraud. Last visited January 27, 2019.

structural and longstanding disability backlogs set the conditions for this massive fraud? Why did it take so long for the agency to detect any patterns? And how can the public be certain that this type of incident will never happen again at SSA?

The Root Causes behind Huntington

At its most basic level, Huntington was a failure of program design. To call the quality assurance system misinformed or weak would be generous – it was virtually non-existent. The entire hearing office apparatus was focused on case output. This is peculiar, as most large operational entities balance three critical elements with exacting detail: output, quality, and speed. Unfortunately, in Huntington and indeed the entire national hearing operation in 2010, there were no quality metrics or data sources for ALJs who paid large numbers of claims nor were their cases flagged for special scrutiny by the agency. In fact, at the time, there was a commonly-held but incorrect belief among SSA managers and executives that any additional scrutiny by management would violate the Administrative Procedures Act (APA). Absent an effective counterbalance of quality assurance, the decades-long backlogs had fostered a one-dimensional culture – reduce the backlog at all costs.

Although it may seem odd today, ALJs who cleared vast number of cases back in 2010 were often celebrated, with insufficient regard about whether they followed relevant rules and regulations when paying claimants. Backlogged cases from around the country would be shipped to these "super-producers" who rushed these cases through, with some approaching 2,000 cases clearances annually.

Amazingly, the Huntington hearing office and staff received accolades and awards for clearing exceptionally high numbers of disability cases. ALJ Daugherty was a driving force in putting Huntington, WV on the map as a highly productive office. The Philadelphia Regional Office, tasked with overseeing Huntington, frequently touted Huntington as a model office. They even recommended the Huntington Hearing Office Chief Judge for a coveted slot as a national trainer of new ALJs. Years later, the Hearing Office Chief Judge would plead guilty to illegal surveillance of a whistleblower in the Huntington office.[94]

In 2005, the agency implemented a new case processing system for the hearing operation known as the Case Processing Management System (CPMS). Wisely, the system built in controls to ensure random case assignment to all ALJs in any office. In other words, in a ten-ALJ office, each ALJ was supposed to receive every tenth case. This was consistent with the random assignment provisions of the APA. Unfortunately, at the national level, this critical anti-fraud control was turned off, allegedly because it was a drag on productivity. This was the loophole that attorney Conn and ALJ Daugherty exploited maximally.

Working together, Daugherty and Conn engaged in an extensive scheme to ensure that Conn's cases were not assigned randomly. Indeed, a significant portion of Conn's cases would wind up assigned to a remote

---

94 The former Social Security Administration Chief Administrative Law Judge for Huntington, WV, Charlie Paul Andrus, pleaded guilty to conspiring to retaliate against a former SSA employee who provided information to investigators. See "Former Chief Administrative Law Judge Pleads Guilty to Conspiracy to Retaliate Against Informant," DOJ Press Release 16-682, June 13, 2016, available at https://www.justice.gov/opa/pr/former-chief-administrative-law-judge-pleads-guilty-conspiracy-retaliate-against-informant. Last visited January 31, 2019.

Kentucky office. Under the scheme, Daugherty would preside over video hearings from Huntington, West Virginia with Conn sitting alone in the Pike, Kentucky remote hearing site. Many of these hearings were not true hearings either – rather, they were little more than a five-minute sham that gave the appearance of a real hearing. In short, undercutting the random case assignment methodology allowed two corrupt players to work together at a remote location via video with no witnesses. Controls are only effective when left in place.

A quick review of case assignment data by local management would have easily detected these anomalies. Indeed, OHO headquarters staff used off-the-shelf social networking software to map work-related inter-relationships between private representatives and ALJs in the Huntington office. It showed glaring irregularities in the case assignment process, including a complete breakdown in the random case assignment process. This chart was a key element in the agency's referral to law enforcement.

The Investigation and Prosecution

Complex fraud cases consume huge blocks of resources. They are difficult to prove, require painstaking detailed investigative work, move glacially, and rarely result in the numeric quick credit that straightforward SSN fraud cases yield. Further, complex fraud cases involving claimant's lawyers bring additional levels of scrutiny and process, invoking special DOJ regulations. Defendants from professional disciplines often come with high-powered defense counsel. For example, in this case, Conn hired Washington insider Abbe Lowell to defend him. The agency owes a debt of gratitude to the tenacious SSA team that fought through prosecutorial resistance and produced vast stacks of evidence to bring this case to justice.[95]

The Clean-up Phase – Better Controls for the Future

With any large entitlement program, nothing erodes trust faster than a massive fraud case. That was both the danger and opportunity presented by the Huntington situation. Do nothing and watch your program wither away. Choose action and you have a tremendous opportunity for significant change management activities that might have ordinarily eluded you under the typical status quo détente of a large Federal program with multiple stakeholders and oversight committees. Seizing the initiative, the agency instituted numerous quality controls to prevent similar cases in the future.[96]

---

95 In October 2018, the President's Council on Integrity and Efficiency recognized several critical players who helped to crack this complex conspiracy case, the Eric Conn Conspiracy Investigation Team. The award was "[i]n recognition of efforts related to the successful multi-agency criminal investigation and prosecution of multiple defendants in the Eric Conn conspiracy case, which uncovered the largest known fraud scheme ever perpetrated against SSA." For a list of the award recipients, see https://www.ignet.gov/sites/default/files/files/21st_Annual_Awards_Ceremony_Program_Web.pdf. Last visited February 1, 2019.

96 These controls include the following:
   A. June 2011 – Systems change so only managers, master docket clerks, and lead legal assistants can assign cases to ALJs.
   B. June 2011 – Improved monitoring of time and attendance at the Huntington office.
   C. August 2011 – Nationwide training for all managers regarding time and attendance.
   D. October 2011 – Hearing Office Chief ALJ removed from his management post.
   E. October 2011 – Begin quarterly nationwide training of all staff on policy issues
   F. October 2011 – Reminder to management judges to rotate medical and vocational experts

This massive fraud case should act as a fraud deterrent for decades to come. Hopefully, it will drive generations of SSA operational leaders to: (1) collect, visualize, and review hearing office data, (2) ensure that fraud controls remain in place, even when they may slow operations somewhat; and (3) build realistic and reliable quality metrics into all workloads at the program design stage.

About the McCrery-Pomeroy SSDI Solutions Initiative

The McCrery-Pomeroy SSDI Solutions Initiative is a project dedicated to identifying practical policy changes to improve the Social Security Disability Insurance (SSDI) program and other policies for people with disabilities. Launched in 2014, the initiative originally commissioned a number of accomplished policy experts from a variety of backgrounds to put forward 12 different policy proposals, each addressing a unique issue with current disability policy. These papers were peer-reviewed, presented at the Initiative's 2015 SSDI Solutions Conference, and ultimately published in the 2016 book SSDI Solutions: Ideas to Strengthen the Social Security Disability Insurance Program. The Initiative's work helped to elevate SSDI to the attention of policymakers and has led to the proposal, enactment, and implementation of numerous legislative and regulatory improvements.

Beginning in 2018, the SSDI Solutions Initiative commissioned seven additional papers designed to build upon the work of the 2016 book. These papers will present additional research, offer implementation guidance, or offer new ideas to further improve disability policy in the United States.

The SSDI Solutions Initiative is co-chaired by former Congressmen Earl Pomeroy (D-ND) and Jim McCrery (R-LA), both former Chairmen of the House Ways & Means Social Security Subcommittee. The SSDI Solutions Initiative is a project of the Fiscal Institute at the Committee for a Responsible Federal Budget.

More information about the SSDI Solutions Initiative is available at http://www.SSDISolutions.org/.

Other Papers in This Series

- "Developing Social Security Disability (SSDI) Reform Demonstrations to Improve Opportunities and Outcomes Based on Lessons Learned" by Jason Fichtner and Jason Seligman

- "Development of an Employment/Eligibility Services (EES) System" by David Stapleton, Yonatan Ben-Shalom, and David Mann

- "Compression and Rationalization: Demarcating the Roles of DDS and ALJs in the Disability Determination Process" by Harold J. Krent

- "Targeting Early Intervention Based on Health Care Utilization of SSDI Beneficiaries by State, with Emphasis on Mental Disorder and Substance Abuse" by Joyce Manchester

- "Implementing a Community-Focused Health & Work Service (HWS)" by Jennifer Christian, Thomas Wickizer, and Kim Burton