

# Social Security Disability Hearings: A Case Study in Quality Assurance and Due Process

Deborah A. Chassman

Howard Rolston

Follow this and additional works at: <http://scholarship.law.cornell.edu/clr>

 Part of the [Law Commons](#)

---

### Recommended Citation

Deborah A. Chassman and Howard Rolston, *Social Security Disability Hearings: A Case Study in Quality Assurance and Due Process*, 65 Cornell L. Rev. 801 (1980)

Available at: <http://scholarship.law.cornell.edu/clr/vol65/iss5/7>

This Article is brought to you for free and open access by the Journals at Scholarship@Cornell Law: A Digital Repository. It has been accepted for inclusion in Cornell Law Review by an authorized administrator of Scholarship@Cornell Law: A Digital Repository. For more information, please contact [jmp8@cornell.edu](mailto:jmp8@cornell.edu).

# SOCIAL SECURITY DISABILITY HEARINGS: A CASE STUDY IN QUALITY ASSURANCE AND DUE PROCESS\*

*Deborah A. Chassman*† and *Howard Rolston*‡

In an article entitled *The Management Side of Due Process*,<sup>1</sup> Professor Jerry L. Mashaw argues that the procedural safeguards typically associated with due process in trial suits are insufficient to ensure fairness in the adjudication of social welfare claims. Social welfare programs, he contends, involve purposes, methods, and clientele that seriously compromise traditional elements of due process, such as the right to notice of adverse claims, the right to present testimony, and the right to cross-examine adverse witnesses. Arguing that additional safeguards are necessary, Mashaw concludes that managers of the social welfare adjudication process could enhance procedural fairness by establishing quality control systems.

According to Mashaw, these inadequacies arise from the basic, widespread assumption that traditional procedural safeguards are "self-correcting mechanisms" that ensure "accurate findings of facts and the authoritative application of law to fact."<sup>2</sup> Consequently, the traditional "attributes of trial-type procedure . . . have come to define due process," and as a result "[c]laims of unfairness are made in terms of the denial of one or

---

\* Although much of this Article derives from the authors' experience in developing and operating a quality assurance system at the Office of Hearings and Appeals of the Social Security Administration, they have no current connection with any such operation. Thus, the information contained herein should not be taken to be representative of any quality assurance system which may currently be operating in OHA, and their opinions and judgments should in no way be taken as indicative of any Social Security Administration opinions or judgments. No official support or endorsement by the Office of Hearings and Appeals of the Social Security Administration is intended or should be inferred.

The authors express their thanks to Professor Jerry Mashaw of Yale University and Professors Roy Shotland and Warren Schwartz of Georgetown University for their encouragement and their helpful comments on earlier drafts.

† B.A. University of Chicago, 1958; M.A. C.U.N.Y., 1971.

‡ B.A. University of California, 1962; M.A. University of Michigan, 1963; Ph.D. Harvard University, 1971.

<sup>1</sup> Mashaw, *The Management Side of Due Process: Some Theoretical and Litigation Notes on the Assurance of Accuracy, Fairness and Timeliness in the Adjudication of Social Welfare Claims*, 59 CORNELL L. REV. 772 (1974).

<sup>2</sup> *Id.* at 775.

more of the[se] attributes.”<sup>3</sup> Mashaw challenges the underlying assumption with evidence suggesting that these traditional elements of due process do not tend to assure accuracy of results in the social welfare setting. From this he concludes that “[d]ue process in the social welfare context . . . requires redefinition to include management processes which will tend to assure the accuracy of claims adjudications.”<sup>4</sup>

Mashaw does not view such management efforts as a replacement for traditional safeguards, but rather as a necessary supplement to them.<sup>5</sup> Mashaw bases his reasoning on three assertions. First, traditional methods of assuring due process are inadequate in the social welfare context. Second, appropriately applied quality control systems can repair this inadequacy, enhancing the “fundamental fairness” of adjudication by increasing its tendency to produce accurate results. Third, quality control systems can, in fact, be successfully applied in the social welfare adjudication setting.

This Article will evaluate the two latter assertions. Although the discussion will not decisively resolve the correctness of Professor Mashaw’s position, it does provide evidence pertinent to such a resolution. Specifically, it will show that although theoretical and practical problems arise in the application of quality control systems to certain forms of social welfare adjudication, such problems are surmountable. Furthermore, current data on the application of quality assurance to social welfare adjudication strongly suggest that such procedures can improve the accuracy of these adjudications.

Evidence supporting these conclusions derives from the authors’ experience in designing, implementing, and operating a quality assurance system at the Office of Hearings and Appeals (OHA) within the Social Security Administration (SSA). OHA is responsible for conducting hearings and issuing decisions on claims for Social Security benefits which the agency has denied at lower levels.<sup>6</sup> The OHA hearing process is relevant to Mashaw’s claim because it involves many of the procedural safeguards cus-

---

<sup>3</sup> *Id.*

<sup>4</sup> *Id.*

<sup>5</sup> *Id.* at 775, 804.

<sup>6</sup> Although OHA is responsible for all Social Security hearings, this Article is concerned only with disability hearings, which have constituted nearly 90% of the OHA caseload. SUBCOMM. ON SOCIAL SECURITY OF HOUSE COMM. ON WAYS AND MEANS, 94TH CONG., 1ST SESS., RECENT STUDIES RELEVANT TO THE DISABILITY HEARINGS AND APPEALS CRISIS 3 (Comm. Print 1975).

tomarily associated with due process, such as the rights to adequate notice, representation, examination, and presentation of evidence. Nonetheless, this Article will show that application of a quality assurance system can enhance the accuracy of these adjudications.

## 1

## THE DISABILITY ADJUDICATION PROCESS

Title II of the Social Security Act provides for the payment of disability insurance benefits to persons meeting insured status who are unable

to engage in any substantial gainful activity by reason of any medically determinable physical or mental impairment which can be expected to result in death or which has lasted or can be expected to last for a continuous period of not less than 12 months . . . .

[A person shall be] determined to be under a disability only if his physical or mental impairment or impairments are of such severity that he is not only unable to do his previous work but cannot, considering his age, education, and work experience, engage in any other kind of substantial gainful work which exists in the national economy. . . .<sup>7</sup>

The Secretary of HEW has promulgated criteria to assist in making disability determinations. According to the regulations, a person is disabled if, in the absence of evidence to the contrary, he suffers from one of many listed impairments or their equivalents.<sup>8</sup> Even if a person's condition is not included in the listed impairments or their medical equivalents, he can still qualify for disability benefits if his condition in fact prevents him from engaging in any substantial gainful activity.<sup>9</sup> To determine whether an applicant satisfies this last criterion, the regulations require consideration of factors including the person's age, education, and work experience.<sup>10</sup>

This last type of determination, often referred to as the residual functional capacity criterion, has probably caused the most

---

<sup>7</sup> 42 U.S.C. §§ 423(d)(1)(A), 423(d)(2) (1976).

<sup>8</sup> 20 C.F.R. §§ 404.1502-1504 Subpart P, App. I (1979).

<sup>9</sup> 20 C.F.R. § 404.1502 (b) (1979).

<sup>10</sup> 20 C.F.R. §§ 404.1502 (b), 404.1504 (b), (c) (1979). Additionally, § 404.1502(c) allows a finding of disability in certain cases where vocational factors are extremely severe.

problems and variations in disability adjudications. Decisionmakers can, to a large degree, use objective criteria to determine whether a person's disability meets or equals the medical listings. The decisionmaker must exercise considerably greater judgment, however, when ruling whether a person whose disability is less severe than the listings still has sufficient residual functional capacity, considering his age, education, and work experience, to allow him to perform a job. This decision requires an elaborate balancing of the severity of the individual's disability with his educational and vocational abilities. These subjective determinations comprise a significant portion of the disability caseload; for example, 23.9% of all disability claims initially allowed in fiscal 1977 were based on the residual functional criterion.<sup>11</sup> With each successive stage of appeal, a higher percentage of claims involves the residual functional criterion. Only about half of all disability claims are eventually decided on the basis of the medical listings or equivalency criteria.<sup>12</sup>

---

HEW, in an attempt to standardize disability determinations involving a mixture of medical and vocational factors, developed new medical-vocational regulations, commonly referred to as the "vocational grid." See 20 C.F.R. Subpart P, App. 2 (1979). This grid directs a conclusion concerning the applicant's disability according to the findings of fact as to a claimant's age, education, work experience and residual functional capacity. For example, the grid directs a determination that a claimant who is limited to sedentary work by his residual functional capacity is disabled when he is age 55 or older, possesses a limited education, and is unskilled. Yet the grid also specifies that a similar individual age 45-49 is not disabled.

The authors are unaware of the grid's success as measured by quality assurance data because the data on which this Article is based were collected before HEW implemented the grid. The grid's ability to reduce ambiguity in disability determinations, however, is limited by its inability to direct a finding of nondisability where mental factors exist, and the presence of such factors is common. Also, use of the grid still requires an administrative law judge's (ALJ) finding of fact on each factor used to calculate the grid requirements and thus can increase uniformity only if the variation in ALJ decisions arises from the combining of factual findings rather than the findings of fact themselves. Although quality assurance errors have not been broken into categories permitting estimation of the percentage of errors that result from the findings of fact rather than from the combining of the findings, major error areas such as insufficient medical evidence and failure to resolve issues of pain would not be reduced because of the grid.

<sup>11</sup> SUBCOMM. ON SOCIAL SECURITY OF HOUSE COMM. ON WAYS AND MEANS, 95TH CONG., 2D SESS., DISABILITY ADJUDICATION STRUCTURE 29-31, Table III (Comm. Print 1978) [hereinafter cited as DISABILITY ADJUDICATION STRUCTURE].

Of the remaining initial allowances, 34.2% were based on a finding that the disability met the medical listings and 41.9% on a finding that the disability was equivalent to the medical listings. *Id.*

<sup>12</sup> M. BERKOWITZ, W. JOHNSON & E. MURPHY, PUBLIC POLICY TOWARD DISABILITY 60 (1976).

The disability program's adjudicatory ladder is unique among Social Security insurance-based programs: it employs a federal-state arrangement for the adjudication of its claims. The Social Security Administration, through its federal district offices, takes claims for disability insurance benefits and sends them to state agencies, which develop medical and other evidence of disability and make determinations of disability.<sup>13</sup> These determinations are then sent to the SSA headquarters (the Office of Disability Operations) where the state determinations are issued as federal decisions.<sup>14</sup> Claimants for disability benefits who disagree with these initial determinations have the right to have their claims reconsidered. Reconsideration procedures follow the same route as initial claims—the SSA district offices receive the requests, the state agency develops the evidence and reconsiders its determinations, and the SSA normally adopts the state determinations.<sup>15</sup> Although the SSA has the power to reduce state agency awards to claimants<sup>16</sup> or return a case to the appropriate state agency for review, it has no authority to issue a determination more favorable than the state agency's. In practice, however, neither of the permitted actions is common.

State agency participation in the process ceases at the third adjudicative level. A claimant still dissatisfied with a reconsideration determination may request a formal hearing before an administrative law judge (ALJ) in the OHA.<sup>17</sup> Requests for hearings are again taken at SSA district offices, but these requests are then forwarded to one of roughly 140 SSA hearing offices located across the nation and operated by the OHA. Requests for hearings are assigned to individual ALJ's who are then responsible for perfecting the evidentiary record, holding face-to-face nonadversary hearings (unless a claimant waives the right to personal appearance), and issuing final decisions.<sup>18</sup> ALJ's can request the

---

<sup>13</sup> See generally STAFF OF SUBCOMM. ON SOCIAL SECURITY OF HOUSE COMM. ON WAYS AND MEANS, 95TH CONG., 1ST SESS., SURVEY OF STATE DISABILITY AGENCIES UNDER SOCIAL SECURITY AND SSI PROGRAMS (Comm. Print 1975).

<sup>14</sup> See 42 U.S.C. § 421(a), (c) (1976); 20 C.F.R. §§ 404.1520 (a), 404.1521 (1979). See generally DISABILITY ADJUDICATION STRUCTURE, *supra* note 11, at 4. See also H.R. REP. NO. 100, 96th Cong., 1st Sess. 3, 8-10, 25-26 (1979).

<sup>15</sup> See 20 C.F.R. §§ 404.909, 404.910 (1979).

<sup>16</sup> See 42 U.S.C. § 421(c) (1976); 20 C.F.R. § 404.1520(c) (1979).

<sup>17</sup> See 42 U.S.C. § 421(d) (1976); 20 C.F.R. § 404.917 (1979).

<sup>18</sup> See 20 C.F.R. §§ 404.922a-404.939 (1979).

appearance of medical and vocational experts at the hearing and can require claimants to undergo consultative examinations. Claimants can elect to be represented by legal counsel or by lay persons, but the hearing remains nonadversarial whether or not claimants are represented.

A claimant dissatisfied with the decision at the hearing level may, within sixty days, request the Appeals Council to review his claim.<sup>19</sup> The Council, a body of attorneys that sits in the Office of Hearings and Appeals headquarters in Washington, D.C., can refuse to review an ALJ decision.<sup>20</sup> The Appeals Council review mechanism is a claimant's last administrative recourse; further appeal must be made through the federal district courts.<sup>21</sup>

Of particular significance in this adjudicatory scheme is the lack of control exercised by the SSA over administration and decisionmaking. Although the federal government funds all benefits and state administrative costs, state law controls many administrative aspects of the program and the SSA routinely accepts state determinations on claims. With state agencies making most disability determinations and ALJ's—whose independence is assured by the Administrative Procedure Act (APA)<sup>22</sup>—making most other decisions, federal control over administration, decisionmaking, and costs is difficult to maintain. Critics of the disability program point to these procedural arrangements, along with the vagueness of the disability definition, as the main problems of the system.<sup>23</sup>

---

<sup>19</sup> See 20 C.F.R. §§ 404.945, 404.946 (1979). For a general description of the hearing and appeal process, see SUBCOMM. ON SOCIAL SECURITY OF HOUSE COMM. ON WAYS AND MEANS, 96TH CONG., 1ST SESS., SOCIAL SECURITY ADMINISTRATIVE LAW JUDGES: SURVEY AND ISSUE PAPER (Comm. Print 1979) [hereinafter cited as ALJ SURVEY]; STAFF OF SUBCOMM. ON SOCIAL SECURITY OF HOUSE COMM. ON WAYS AND MEANS, 94TH CONG., 2D SESS., DISABILITY INSURANCE—LEGISLATIVE ISSUE PAPER (Comm. Print 1976); J MASHAW, C. GOETZ, F. GOODMAN, W. SCHWARTZ, R. VERKUIL & M. CARROW, SOCIAL SECURITY HEARINGS AND APPEALS (1978) [hereinafter cited as HEARINGS AND APPEALS].

<sup>20</sup> See 20 C.F.R. § 404.947 (1979). The Appeals Council can also reopen a case on its own motion within sixty days after notice of the ALJ's decision.

<sup>21</sup> See 42 U.S.C. §§ 405(g), 421(d) (1976).

<sup>22</sup> See generally ALJ SURVEY, *supra* note 19, at 12-21; text accompanying notes 32-33, *infra*.

<sup>23</sup> See, e.g., ALJ SURVEY, *supra* note 19, at 14-15. The Disability Insurance Amendments of 1979, H.R. 3236, 96th Cong., 1st Sess. contain provisions intended to alleviate some of these problems. The bill strengthens the federal role in the disability adjudication process by giving the Secretary of Health, Education and Welfare the authority to establish procedures and performance standards for state disability determinations. States unwilling or unable to comply with the standards would relinquish direct administration to the federal government.

The bill would also require the SSA to review 65% of all state agency allowances before implementation of the determinations and payment of benefits. This review would

Both the division of decisionmaking responsibility and the subjective nature of many disability determinations contribute to large interstate variations in claim approval rates, frequency of error rates, processing times, and case processing costs. For fiscal year 1977, state agency denial rates ranged from 40.2% to 66.7%.<sup>24</sup> State error rates, based on a sample review for the period May to October, 1977, ranged from 6% to 27.7%.<sup>25</sup> Processing times varied widely, and state agency productivity ranged from 156 to 310 cases processed per employee.<sup>26</sup> Similar variations exist in ALJ approval rates: some ALJ's awarded benefits in only about 10% of their cases while others awarded benefits in almost 90% of their cases.<sup>27</sup>

In the mid-1970's a crisis arose in the disability program. In addition to the variations among state agencies and ALJ's, the number of initial claims and the number and proportion of cases appealed up the administrative ladder and taken eventually to the federal courts dramatically increased. In 1968, 317,400 (47.7%) of the 638,900 disability claims brought were allowed on initial determination. Over 67,000 claimants (21.0% of all claims initially denied) sought reconsideration and almost 35% of these claims were allowed upon reconsideration. Approximately 20,200 claimants (45.9% of all reconsiderations lost) requested a hearing and ALJ's or the Appeals Council allowed claims in about 42% of these. That same year, 916 claims (3.8% of all hearings lost) were taken to court and 313 of these were eventually allowed.<sup>28</sup> By fiscal year 1977 both the number of claims and the percentage allowed at each level had changed drastically. The number of claims brought exceeded 1,250,000, but only 38.4% were allowed initially. Almost 220,000 disappointed claimants (28.4% of all claims initially denied) sought reconsideration and only 27.7% of these were allowed on reconsideration. Over 92,000 claimants (58.0% of all reconsiderations lost) sought a hearing and 48% of these claims were allowed by ALJ's and the Appeals Council. Finally, claimants filed almost 5,900 court actions (12.1% of all

---

be phased in with a 15% review in 1980 and the full 65% review by 1982. The bill was passed by the House of Representatives on September 6, 1979. See H.R. REP. NO. 100, 96th Cong., 1st Sess. 3, 8-10, 25-26 (1979).

<sup>24</sup> DISABILITY ADJUDICATION STRUCTURE, *supra* note 11, at 28, Table 11.

<sup>25</sup> *Id.* at 42-43, Table XI.

<sup>26</sup> *Id.* at 41, Table X.

<sup>27</sup> HEARINGS AND APPEALS, *supra* note 19, at 21 & Fig. 1-2, 42-43.

<sup>28</sup> DISABILITY ADJUDICATION STRUCTURE, *supra* note 11, at 24-27, Table 1.

hearings lost) and were successful in about 22% of the cases decided.<sup>29</sup> The percentage of claims allowed at the initial determination compared with the total number of claims ultimately allowed fell from 90.9% to 81.9% over a decade.<sup>30</sup>

This combination of trends—the huge increase in the number of disability claims filed, the decreasing percentage of claims allowed by the state agencies at the initial and reconsideration levels, the increase in claims allowed at the hearing and appeal levels, the continuing diversity in state procedures, and increased error rates—raised serious questions concerning the efficiency and efficacy of the system. These problems were particularly evident at the hearing level, where ALJ's allowed a high percentage of claims twice-denied by state agencies, deciding many of these claims under the subjective residual functional capacity criterion. Exacerbating matters, the rapid increase in requests for hearings had caused delays, frequently up to a year,<sup>31</sup> in the processing of cases at the hearing and appeal levels.

## II

### THE OHA SETTING

OHA implemented its quality assurance system in the context of increased workload pressure, aroused skepticism, and stricter scrutiny. The particular circumstances of OHA adjudication dictated in large measure the form of its quality assurance system and the issues that arose regarding its applicability.

One of the most important aspects of OHA adjudication is the independence of the ALJ's under the APA. The purpose of this independence is to guarantee claimants the right to an independent *de novo* decision, neither biased by prior agency determinations nor affected by agency pressures on the decisionmaker. Various procedures exist to protect ALJ's from agency interference. For example, ALJ's are hired by the Office of Personnel Management, not by the particular agency, and are not subject to annual agency evaluation as are other Civil Service employees.

In addition to limitations in the employment relationship between the agency and ALJ's, procedural safeguards preserve the independence of ALJ's. Prior to an ALJ's decision on a claim, for

---

<sup>29</sup> *Id.*

<sup>30</sup> *Id.*

<sup>31</sup> See ALJ SURVEY, *supra* note 19, at 21.

example, *ex parte* contact between the ALJ and the agency on the substance of the claim is proscribed.

The independence guaranteed ALJ's by the APA does not bar the agency from monitoring various aspects of the performance of an ALJ.<sup>32</sup> The freedom from agency interference refers to decisional interference only and is designed to protect the claimant, not the ALJ's. Nonetheless, a large number of OHA ALJ's had a much broader view of their independence under the APA; they objected vehemently to OHA monitoring or controlling any aspect of their performance, including management of staff resources. OHA, meanwhile, recognized its dependence on the good will of ALJ's to hear the cases and reduce the pressing backlog. In short, OHA walked on eggs as it instituted the quality assurance system. Even though it could have legally exerted greater control and more closely monitored ALJ processing of claims, OHA designed the quality assurance system to avoid, as much as possible, directly assaulting the view of the ALJ's of their independence.<sup>33</sup>

In addition, since quality assurance was a new system imposed on existing agency processes, OHA tailored it to meld with the ALJ hearing process. In order to evaluate the OHA quality review system, it is therefore necessary to describe normal OHA and SSA processing procedures for ALJ hearing decisions.

When an ALJ makes an unfavorable determination on a claim, the decision is sent to the claimant, and the claims file (the evidentiary and procedural record) is forwarded to the OHA central office; the file remains available in the event a claimant requests review by the Appeals Council. If a favorable decision is issued, however, the file is sent to the Office of Disability Operations of SSA for payment of any past-due benefits, preparation

---

<sup>32</sup> For an excellent discussion, including the opinions of several legal scholars, on the application of the Administrative Procedure Act to Social Security cases and the independence issue, see ALJ SURVEY, *supra* note 19, at 6-11.

<sup>33</sup> ALJ independence has been an issue in a number of recent civil suits brought by ALJ's and claimants against the SSA. One such suit, *Bono v. Califano*, No. 77-0819-CU-W-4 (W.D. Mo., settled June 7, 1979), was settled. Although the provisions of the settlement affirm the authority of the SSA and OHA to exercise administrative and management functions over their corps of ALJ's, it does not appear to deal with possible future improvements in the quality review system. In addition, the settlement jeopardizes the feedback potential of Appeals Council review by providing that cases remanded by the Council will be assigned to an ALJ different from the one who originally decided the case. For a discussion of the issues in the *Bono* case and related suits, see ALJ SURVEY, *supra* note 19, at 112-13 (App.J).

for regular computer payment of future benefits, and various other less important processes, all of which typically require about three months to complete. The significant aspect of this procedure is that the claims file is unavailable to OHA while it is being processed for payment. Thus, if the Appeals Council wishes to review a *favorable* ALJ decision, it must review it either before the payment process begins or else wait three months until the Office of Disability Operations has completed all payment actions.<sup>34</sup>

When, within the sixty day time period for appeal, a claimant requests that an *unfavorable* ALJ decision be reviewed by the Council, the case file is available and the review can be conducted immediately. The Council's support staff—over 200 hearings and appeals analysts—initially reviews the file. This staff is necessary because of the large volume of appeals; it would be impossible for the Council members to review personally over 4,000 appeals per month (roughly 285 appeals per Council member) and also meet the additional workload generated by further appeals to the district courts.<sup>35</sup> The analysts are highly trained in the relevant aspects of Social Security programs and are assigned to the work of only one or two Council members. The analysts do not technically make the decisions for the Council, but instead recommend that the Council either deny or grant review. These pre-existing case processing procedures significantly influenced the design of the quality assurance adopted by OHA.

### III

#### QUALITY ASSURANCE SYSTEMS

Mashaw asserts that "a quality control system can be adapted to virtually any type of enterprise or end product for it involves merely the development of standards, the evaluation of performance against those standards, and action to upgrade substandard performance."<sup>36</sup> The first step, the development of standards, essentially involves the formulation of an error definition which can be applied with some precision to divide the "end products"

---

<sup>34</sup> See text accompanying note 42 *infra*.

<sup>35</sup> The Appeals Council processed 52,273 requests for review in 1977, or almost 4,400 a month. SOCIAL SECURITY ADMINISTRATION, BUREAU OF HEARINGS AND APPEALS, BHA FACT SHEET FOR FISCAL YEAR 1977, BHA PUB. NO. 039 (1978) [hereinafter cited as BHA FACT SHEET]. The Bureau of Hearings and Appeals (BHA) is now the Office of Hearing and Appeals (OHA).

<sup>36</sup> Mashaw, *supra* note 1, at 791.

into defective and nondefective categories. Evaluation of performance against those standards, the second stage, requires the establishment of various procedures for applying the error definition and detecting any errors. The final step, action to upgrade substandard performance, includes both correction of errors discovered by the error detection process and analysis of information acquired through that process so that management can prevent future errors. This Article will discuss these three stages in sequence. As the discussion will show, however, these three components are interwoven and cannot be treated entirely separately.

#### A. *Error Definition*

The OHA quality assurance system defined an error as any defect so serious that it either required corrective action by the Appeals Council, or would have required such action had the claimant requested Council review within the sixty-day limit.<sup>37</sup> This definition is significant for two reasons. First, it mandates that only "sufficiently serious" defects are errors. Second, it envisions that a subsequent operational level, the Appeals Council, will apply this standard to define errors made at the lower hearing level.

##### 1. *Sufficiently Serious*

The main reason for restricting the error definition to serious defects was to focus the system on the most egregious problems. The sheer magnitude of the disability hearing process and the time pressures under which ALJ's work commonly cause minor deviations from the theoretical norm. Another reason for limiting the error definition to serious inadequacies is that, given the subjectivity of disability determinations and the negative attitude of many ALJ's towards the quality assurance undertaking, it appeared that the system would be better received and have greater credibility if it identified only serious and relatively clear-cut defects.<sup>38</sup>

---

<sup>37</sup> See note 19 and accompanying text *supra*.

<sup>38</sup> See note 41 and accompanying text *infra*.

## 2. Operational Standard

The judgmental nature of ALJ decisions and OHA's concern with credibility were also the reasons behind the adoption of an operational standard to define error. The subjectivity inherent in disability adjudication stands in sharp contrast to the factors which shape quality assurance systems in the industrial context. For example, a factory that produces sheets of glass can test the tensile strength of the glass by applying a given longitudinal stress to a random sample of glass; it either breaks or it does not. In this setting, quality assurance standards are well-defined. Disability criteria, on the other hand, are not so easily defined, especially where the medical listings or equivalents do not apply.<sup>39</sup> Although subjectivity pervades the entire disability determination process, it is typically most troublesome in borderline cases, which most often reach OHA.

These considerations suggest that the question of what constitutes an error is not so much the issue of what abstract definition to apply, but rather the problem of who shall apply whatever definition is adopted. The ability of a quality control system to measure performance authoritatively depends ultimately on the legitimacy of the body that determines which defects shall count as errors. Within OHA, only the Appeals Council has the authority to label an ALJ's decision an error. Consequently, the operational practice of the Appeals Council was adopted as the standard for evaluating ALJ performance.

The use of the Appeals Council's operational standard to define and detect errors in the handling of claims by ALJ's created some troubles. One major problem resulted from the substantial evidence rule,<sup>40</sup> which states that the Appeals Council will not disturb an ALJ's decision that is based on substantial evidence. This rule demands more than a scintilla but less than a preponderance of evidence.<sup>41</sup> While an ALJ's decision is incorrect if not supported by a preponderance of the evidence, under the operational definition of error that decision is not an error unless it also fails to meet the substantial evidence test. Although this reduces the

---

<sup>39</sup> See notes 8-10 and accompanying text *supra*.

<sup>40</sup> 20 C.F.R. § 404.947a(a)(3) (1979).

<sup>41</sup> In *Consolo v. Federal Maritime Comm'n*, 383 U.S. 607 (1966), the Court defined substantial evidence as "something less than the weight of the evidence, and the possibility of drawing two inconsistent conclusions from the evidence does not prevent an administrative agency's finding from being supported by substantial evidence." *Id.* at 620.

scope of the quality control system, it only eliminates marginal errors, automatically focusing attention on the more serious defects.

The operational standard also raises questions of uniformity in its application. For example, when the workload is heavy, the Appeals Council may tend to construe substantial evidence liberally and affirm an ALJ's decision which is supported by less evidence than the Council would require during slack periods when it had ample time to scrutinize each case. If such shifts occur, then error rates would change artificially.

### B. *Error Detection*

The hearings and appeals analysts on the Appeals Council staff actually perform the quality assurance case review. A random sample of approximately five percent of all ALJ decisions is computer-selected roughly fifteen days after the close of the sample month. If the claimant has appealed the decision, the Council conducts the quality assurance analysis as part of the required review. Unappealed sample denial cases are typically reviewed by the analysts after the sixty-day period for filing an appeal has passed. The analysts review the claims files for the sample and complete a form noting claimant's age, education, and work experience, and information concerning the processing of the claim, including the date of the hearing, a summary of any expert testimony, and an evaluation of any defects, serious or nonserious, that are present. If the analyst detects a serious error, he forwards the claims folder and completed form to the Appeals Council for concurrence. The case is deemed an error for quality review purposes only if the Council member agrees that a serious defect exists. The Appeals Council never reviews unappealed cases that the analyst decides are error-free. Following the review process, quality review forms are sent to the Division of Appraisal—which is responsible for the overall administration of OHA's quality assurance system—for compilation and analysis.

Quality assurance review of allowance decisions follows a similar path but only after payment and other effectuation actions have been completed.<sup>42</sup> One upshot of adopting a post-effectuation review system for ALJ allowances is that the sixty-day

---

<sup>42</sup> See text accompanying note 34 *supra*.

period during which the Appeals Council may correct an erroneous decision on its own motion has run before the Council reviews these cases.<sup>43</sup> Consequently, the Council and the analysts may treat the allowance decision less seriously. The extent to which this inability to correct wrong decisions affects the analysts' motivation is a matter of conjecture. To the extent that it creates an attitude of futility, the analysts are likely to report fewer errors, especially since the operational definition of error offers considerable leeway. This would systematically understate the error rate.<sup>44</sup>

### C. Error Correction

#### 1. Individual Sample Case Error Correction

Error correction is the key step in the quality assurance system. It should focus not only on correcting errors detected as part of the review process, but also on systematically determining and remedying the causes of errors. Both individual case correction and systematic process correction come within Mashaw's "action to upgrade substandard performance."<sup>45</sup> The proper mix of these error-correction devices depends on the complexity of the process to be reviewed, the amount of time and resources needed for each review, and the level of control the reviewer can directly exercise on the end products. For example, in a manufacturing context, the quality assurance system might consist of an end-of-production-line inspector who tests every manufactured part to determine if it meets the buyer's specifications. Those parts that fail the quality test are discarded. If several rejects are traced to one machine, action will be taken to repair it. The inspector tests each part using clear and objective criteria. He can make rapid decisions without significantly slowing final acceptance of the goods and he can inspect the work of many machines in a relatively short time.

This example posits an inspection system so short, simple, and inexpensive that every product can be examined, all deficient articles rejected, and action easily taken to correct processing de-

---

<sup>43</sup> See note 20 *supra*. The Appeals Council can reopen a case within four years after the date of the notice of the initial determination where there is an error "on the face of the evidence." 20 C.F.R. §§ 404.957(b), 404.958(c) (1979).

<sup>44</sup> See text accompanying note 56 *infra*.

<sup>45</sup> Mashaw, *supra* note 1, at 791.

facts. As the production process becomes more complex or the decision of right or wrong becomes more judgmental, however, the costs of inspecting every final product will increase and the accuracy of the review will probably decrease. Cost-effectiveness is not the only consideration. In OHA, for example, 100% review might be cost-effective given the high average benefit award.<sup>46</sup> But pressure for timeliness in a time-consuming process, limited manpower resources, and the political costs of hiring a staff large enough to review every case may mandate a highly selective review process in which systematic error correction predominates over individual case correction.

The structure of the OHA quality assurance system reflects these considerations. The hearings and appeals analysts review a random sample of only about five percent of all cases, and therefore a maximum of only five percent of all errors can be corrected in conjunction with the review process. In addition, only the appealed denials, which constitute about one-quarter of the sample cases,<sup>47</sup> are subjected to quality assurance review within the sixty-day limit for Council corrective action, reducing the possibilities for individual case correction even further. As alternatives to individual sample case error correction, OHA has attempted to influence ALJ's by returning their erroneous cases, reporting errors and dispensing instructional materials.

## 2. *Feedback and Instruction*

OHA has at times incorporated a remand procedure for erroneously decided ALJ decisions. The predominant practice of the Appeals Council, however, has been to make any necessary corrections itself, unless a new hearing is required. In any event, a

---

<sup>46</sup> The House Ways and Means Committee, recognizing cost-effectiveness potentials, approved a provision in H.R. 3236, 96th Cong., 1st Sess. (1979), mandating federal review of 65% of all state agency allowance decisions by 1982. According to Congressional Budget Office (CBO) cost estimates, the 65% review will result in a net savings. CBO based its estimate on a six month test review of 6,299 initial disability allowances where 23.6% of the allowances were returned to the states and 22.1% were subsequently denied. From these percentages of those returned CBO projected the total number of erroneous allowances that would be denied with the 65% review and concluded that the provision would save \$99 million in 1982. H.R. REP. No. 100, 96th Cong., 1st Sess. 36 (1979). The Report does not discuss what went into the decision to limit the review to 65% and no figures are supplied to suggest that 65% is the most cost-effective level of review.

<sup>47</sup> In 1977, ALJ's processed 186,822 cases and claimants appealed 47,719 of these, or 25.5%, to the Appeals Council. Since the sample is a simple random one, about 25% of all sample cases are appealed denials. See BHA FACT SHEET, *supra* note 35, at 2-3.

remand procedure would have been a mixed blessing in the OHA setting.

First, the Council could potentially remand only the sample cases appealed to it or appealed denials, which constitute only one-quarter of all sample cases.<sup>48</sup> Second, feedback of errors should have instructional value and discourage repetition of past mistakes. But OHA remanding of a slew of cases for correction might meet with a storm of protest from a corps of quasi-independent ALJ's on the ground that judgmental findings rather than errors are involved. This in turn might create an operational tendency to return only the most grievous errors. Third, in remanding a case the Appeals Council does not demand that the ALJ reverse his decision. The remand simply contains specific instructions concerning the development of evidence. Under these circumstances, the remand might force a claimant to come before an unusually harsh ALJ a second time. Finally, since the Council could only remand erroneous appealed denials, ALJ's might tend to allow close cases rather than risk a remand. The focus of Appeals Council review on appealed denials instead of allowances already skews the system in this direction to some extent.

### 3. *Error Analysis and Instruction*

An alternative to feedback by remand devised by OHA is a program for isolating specific causes of error and then instructing the ALJ's to remedy these causes. OHA collects data on all facets of case processing, including considerable amounts of medical, vocational, and demographic data, for each kind of decision-allowance, appealed denial, and unappealed denial. From this information, OHA prepares quarterly quality assurance reports. Each report contains a statistical summary of national and regional error-rate data on time-line graphs, a breakdown of errors by type and stage of process, analytic summaries of high frequency problem areas, and recommendations for improvement. In addition, each report focuses on a different adjudicative factor, disability category, or case type, discussing its possible relationship to error-proneness. These quarterly quality review reports are distributed to top management and to the supervisory judge corps, both in the central office and the regions. OHA also sends instructional material on specific error areas to all ALJ's. The Ap-

---

<sup>48</sup> See note 47 *supra*.

peals Council also addresses such topics in regular written reports and face-to-face instructional sessions with hearing offices.

Instructional bulletins and special training can be effective where errors reflect deficiencies in knowledge, but they only suggest rather than demand change. Furthermore, because such efforts are not immediately connected with the correction of a particular error in a particular case, management and adjudicators consider them less important. The authors believe that as long as the backlog of claims remained high, only limited effort was put into detailed instruction, particularly where the instruction emphasized case processing requirements that lengthened the hearing process.

The statistics also indicate that the instructional and reporting programs had no significant effect on the accuracy of the hearing process. In spite of these programs, error rates fluctuated as much as fifty percent from one quarter to the next.<sup>49</sup> Errors based on insufficient medical or vocational evidence, an area of major instructional concentration, also varied markedly with no trend towards improvement.<sup>50</sup>

#### *D. Error-Prone Analysis and Appeals Council Review of Error-Prone Cases*

There is no hard evidence that the OHA quality assurance system has thus far enhanced due process for disability claimants by increasing the accuracy of the hearing process. What then is left of Mashaw's claim that quality control can work? Data derived from the OHA quality assurance system strongly suggest that a third variety of error correction could, if used, improve the accuracy of the disability determination process.

Quality assurance error analyses indicate that certain categories of cases are more error-prone than others. These analyses isolate characteristics that make the work of ALJ's error-prone, and processing practices that contribute to error. Once these character-

---

<sup>49</sup> Appraisal Staff, Division of Administration, Bureau of Hearings and Appeals, Social Security Administration, Report on Quality Review 8-10, April-June 1977 (on file at the Cornell Law Review) [hereinafter cited as *Quality Review April-June 1977*].

<sup>50</sup> Evidentiary errors were found in 51% of all cases in the third quarter of 1976, 42% in the next quarter, 32% in the next, and 44% in the second quarter of 1977. See Appraisal Staff, Division of Administration, Bureau of Hearings and Appeals, Social Security Administration, Report on Quality Review, January-March 1977, April-June 1977, July-September 1977 (on file at the Cornell Law Review).

istics are identified, the Appeals Council can, on its own motion,<sup>51</sup> scrutinize and correct errors in those cases having such characteristics. This form of error correction has several advantages, particularly in a setting like OHA. Foremost, if sufficiently strong error-indicators can be isolated, a relatively small amount of scarce personnel resources could correct a significant number of errors. Second, OHA could accomplish this without an additional review authority.

OHA has made some error-prone analyses, although it has not undertaken any concerted efforts to implement such a system.<sup>52</sup> When OHA established the quality review system, it collected large amounts of data and subjected them to a variety of error-prone analyses. The results of these analyses preliminarily indicate the form the error-prone review should take.

For example, as mentioned earlier, the claim allowance rates of individual ALJ's vary from ten to ninety percent.<sup>53</sup> Simple statistical tests reveal that the error rate for sample cases of ALJ's with extremely high or extremely low allowance rates is higher than for other cases. This is particularly true for the allowance decisions of ALJ's with high allowance rates.<sup>54</sup> This suggests that in some instances the outcome of a claimant's case depends on the luck of the draw rather than the merits of the claim. Appeals Council review of such cases would enhance the fundamental fairness of the system precisely where it has deteriorated.

The data suggest that appealed denials are more error-prone than unappealed denials. Evidently claimants are able to discriminate between correct and incorrect disallowance decisions and thus the error rate for unappealed denials is lower than the rate for appealed denials.<sup>55</sup> The data may be misleading: the futility

---

<sup>51</sup> See note 20 *supra*.

<sup>52</sup> The Division of Appraisal prepared an issue paper outlining its findings on the use of error-prone analysis. See Appraisal Staff, Division of Administration, Bureau of Hearings and Appeals, Social Security Administration, Comprehensive Review of ALJ Reversals, July 25, 1978 (internal OHA document) [hereinafter cited as Review of ALJ Reversals]. OHA also has taken some preliminary steps toward implementation, including consideration of modifying the OHA computerized management information system to aid in the selection of error-prone cases and making a test run. As far as the authors know, however, OHA has not acted further.

<sup>53</sup> See note 27 and accompanying text, *supra*.

<sup>54</sup> See HEARINGS AND APPEALS, *supra* note 19, at 3-4.

<sup>55</sup> OHA sampled appealed and unappealed denials and discovered that the error rate for unappealed denials is between 10% and 25% of the rate for appealed denials. For instance, the error rate during the second quarter of 1977 for appealed denials was 6.5%, and the rate for unappealed denials was 1.6%. Quality Review April-June 1977, *supra* note 49, at 8-10.

problem may artificially reduce the number of errors detected by the analyst in unappealed denials since errors cannot be corrected.<sup>56</sup> A validation study conducted one year after implementation of the quality assurance system discounted this possibility. In that study, OHA's Division of Appraisal reviewed a subset of quality assurance sample cases. The review was conducted without access to the original quality review form. Where the Division detected an error not previously reported, it forwarded the case to the Appeals Council for concurrence. The study revealed that the analysts left no more errors undetected in unappealed denials than in appealed denials.<sup>57</sup> Thus, the futility problem does not undercut the claimant appeal as a factor in identifying error-prone cases.

The data suggest that allowances also are more error-prone than unappealed denials.<sup>58</sup> Furthermore, the validation study indicated that the OHA quality control system understates error-rates for allowances more than for denials.<sup>59</sup> Thus, since all appealed denials are already reviewed by the Appeals Council, the study shows that allowances rather than unappealed denials would be a more profitable area for error-prone review.

Cases containing many error-prone factors can be detected by discriminant analysis, a statistical tool which can categorize a sample of cases into two or more groups based on statistical properties.<sup>60</sup> OHA conducted a trial discriminant analysis study on hearing decisions in 1978. It derived a discriminate function that assigned numerical values to characteristics in accordance with their correlation to error-proneness. A case's error-proneness was represented by a score equal to the sum of its weighted values for

---

<sup>56</sup> See text accompanying note 44 *supra* (post-effectuation review creates futility problem because Appeals Council can no longer correct erroneous decision).

<sup>57</sup> Appraisal Staff, Division of Administration, Bureau of Hearings and Appeals, Social Security Administration, Report on Quality Review Form Validation Study, November 1, 1977 (internal OHA document) [hereinafter cited as Validation Study]. The study, completed in October 1977, reviewed 73 unappealed and 72 appealed denials. It uncovered two additional errors in each category. See *id.* at 2.

<sup>58</sup> For example, during the second quarter of 1977, the error rate for allowances was 6.5% and the rate for unappealed denials was 1.6%. See Quality Review April-June 1977, *supra* note 49, at 8-10.

<sup>59</sup> In the validation study, allowances contained an additional 4.7% of errors whereas unappealed denials contained an additional 2.8%. See Validation Study, *supra* note 57, at 4.

<sup>60</sup> For a full description of discriminant analysis see N. NIE, C. HULL, J. JENKINS, K. STEINBRENNER & D. BENT, STATISTICAL PACKAGE FOR THE SOCIAL SCIENCES 434-67 (2d ed. 1975).

each relevant characteristic. The study concluded that sufficiently strong predictors of errors can be identified so that, for example, a group containing ten percent of the most error-prone ALJ allowances as selected by discriminant analysis would yield about forty percent of all ALJ allowance errors.<sup>61</sup>

These test results indicate that discriminant analysis could significantly improve the accuracy of the hearing process at a relatively low cost.<sup>62</sup> Discriminant analysis would rank new allowance cases according to their error-prone scores. The Appeals Council could then choose either to scrutinize all allowance cases exceeding a certain error-prone score or to review a set number of the most error-prone allowance cases.<sup>63</sup> For individual case correction, the review must occur before effectuation, since the Council has only sixty days to correct errors on its own motion.<sup>64</sup>

---

<sup>61</sup> See Review of ALJ Reversals, *supra* note 52. The report also included a cost-benefit estimate that the 10% error-prone case review would increase costs about \$.8 million but could save as much as \$24 million. See *id.* at 3.

When the Division of Appraisal proposed that OHA use error-prone review to reduce errors, the concept and the techniques for such procedures were largely untried in social welfare settings. Error-prone analysis has now become a common and successful error reduction device in the Aid to Families With Dependent Children (AFDC) and Supplementary Security Income (SSI) programs, both of which are under the jurisdiction of the SSA. The Office of Family Assistance, the SSA component that oversees the state operations of the AFDC programs, recently invited state administrators to a three-day conference on error-prone profiling. The conference took the form of "how to" seminars, with states that had successfully implemented the system reporting on the techniques they used and the relative advantages of different methodologies in different settings. West Virginia, which has the longest running system, reported that its error-prone profiling reduced payment error-rates from 7% to 5% in six years. South Carolina reduced its payment error-rate from 7.8% to 7.1% in less than one year. The District of Columbia instituted error-prone review at the General Accounting Office's mandate and selected a 20% sample of cases for review that had twice the error-rate of a randomly selected sample. In a demonstration project in New Hampshire involving a portion of the Medicaid workload the profiles increased the error yield of the random sample by 250%. The SSI program also reported that the use of profiling in some of its offices substantially reduced error-rates in comparison with "control" offices. The redetermination error rate for the first nine months of 1978 for offices using profiling procedures was 1.7%, while in the offices using standard procedures it was 3.0%. See U.S. OFFICE OF FAMILY ASSISTANCE, CONFERENCE ON THE UTILIZATION OF CHARACTERISTIC PROFILES AS A WORKLOAD PLANNING TECHNIQUE (1979).

<sup>62</sup> Since the review process is not static, continued random sampling to determine error-prone characteristics would be needed to ensure the accuracy of the discriminant analysis.

<sup>63</sup> The marginal productivity of error-prone review decreases as the sample increases. Nevertheless, error-prone review would always uncover more errors than would a random sampling of the same number of cases.

<sup>64</sup> See text accompanying notes 20, 34, & 42, *supra*. The ranking procedures would delay the effectuation of allowance cases. The delay need not be excessive if the ranking only uses precise, easily-identifiable error-prone characteristics. Personnel in the office

Use of discriminant analysis would permit the Appeals Council to detect errors more efficiently and to correct these errors by individual case correction. Process correction efforts could still continue but they would no longer be the primary tool of error correction. With discriminant analysis, the Council could additionally review the most error-prone cases within the most error-prone class of cases—allowances. Appealed denials, the other class of error-prone cases, are already reviewed, and unappealed denials, which are relatively error-free, would not be reviewed. Such an efficient use of Council resources could potentially correct thirty to forty percent of all errors,<sup>65</sup> significantly increasing the accuracy of the disability hearing system.<sup>66</sup>

#### CONCLUSION

Data derived from the OHA experience support Mashaw's claim that quality assurance can be applied to the social welfare adjudication process, even where the decisions involved are highly judgmental. Although the nature of the disability hearing process limits the capability of any review process to correct *all* errors, quality assurance can correct a significant number of defects. Another limitation is the unreliability of error rates measured on a temporal dimension; rate changes may result from changes in what is measured—ALJ behavior—or changes in the measuring instrument—the Appeals Council. Nonetheless, these limitations point to the necessity of careful use of the data derived from OHA's quality assurance system, not to the futility of the system. Even though there is no evidence that the quality assurance system has improved the "fundamental fairness" of the disability hearing process, data derived from the system suggest that use of discriminant analysis could do so.

This conclusion is paradoxical in light of Mashaw's original goal. Mashaw envisioned quality control as a management tool to

---

where all allowance files are sent for effectuation could perform the task. Effectuation of the highly error-prone allowance cases selected for quality assurance review would of course be delayed further pending that review, but not beyond the 60-day limit.

<sup>65</sup> See Validation Study, *supra* note 57, at 2.

<sup>66</sup> Use of discriminant analysis would increase the number of review cases because the current 5% random sampling would still be necessary to update the error-prone characteristics upon which the analysis relies. The review of the error-prone cases would be in addition to the number of cases currently reviewed. However, the savings from the error detection yield from discriminant analysis more than justifies this added cost. See Review of ALJ Reversals, *supra* note 52.

ensure due process so that a claimant could obtain the benefits to which he is entitled. This Article suggests, however, that the primary target for error correction should be allowance cases, entitling fewer rather than more claimants to benefits. Although this may be the immediate effect, in the long run elimination of erroneous allowances could aid those entitled to benefits by reducing pressure to cut disability payment rates or coverage.<sup>67</sup>

---

<sup>67</sup> H.R. 3236, 96th Cong., 1st Sess. (1979), in addition to the provisions for greater federal control over disability claims, would also limit maximum disability benefits and change the mix of disability benefits away from younger disabled workers. *See* H.R. REP. No. 100, 96th Cong., 1st Sess. 2-3 (1979).