

**Address of  
Sylvester J. Schieber, Chairman  
Social Security Advisory Board to the  
2007 National Educational Conference  
Association of Administrative Law Judges  
July 26, 2007**

Good morning. I appreciate the opportunity to talk with you about the state of the Social Security program today from my perspective as a member and chairman of the Social Security Advisory Board.

Nine and a half years ago, in January 1998, I was appointed to the Social Security Advisory Board. In August of that year, the Board issued the first of several reports concerning the disability programs. In that report, the Board said:

Today, as in the past, there are serious concerns about the lack of consistency in decision making; unexplained changes in application and allowance rates; the complexity, slowness and cost of the application and appeals process; the lack of confidence in the system; and the fact that few beneficiaries are successfully rehabilitated so they can become part of the economic mainstream.

I wish I could look back over the 9 years since those words were written and say that things have greatly improved. Unfortunately, the facts surrounding the program would not support such a statement.

There have indeed been some changes for the better.

In particular, the increased uses of technology such as video conferencing, digital recording, and, of course, the electronic folder. Bringing these advances to bear on the adjudication process is already paying dividends. And it will continue to do so. Any major technology initiative goes through a period of start-up problems that need to be resolved. Users need to gain familiarity with the new systems, capacity issues have to be addressed, and program glitches have to be corrected. But an electronic processing environment provides huge opportunities to improve efficiency

and to facilitate the kind of analysis that is essential to resolving longstanding quality and consistency issues.

But, for the most part, the problems that the Board cited in 1998 are largely still problems. Unexplained and substantial inconsistency geographically, among components, and among adjudicators within components remains. The implementation of the Ticket to Work program has not changed the fact that achievements in the area of return to work are almost non-existent. Confidence in the program, to the extent that it can be measured by such things as news articles and Congressional hearings, remains low. And, as you all are painfully aware, the part of the program that seems to be in the worst shape of all is the hearing process.

The number of people who have applied for a hearing but have not yet got a decision is at a historically high level of just under 750,000. That is more than double the 334,000 pending claims at the time in 1998 when the Advisory Board issued its first report on problems in the disability program.

Of course, more cases have been coming in the front door in recent years, but that doesn't come close to explaining the worsening pending levels. Over the past 20 years, hearing office pending levels have been typically around 50 to 80 percent of receipts. In 1996, there was a spike to a then historic high of 95 percent of receipts after which they declined to 63 percent in the year 2000. Since then, they have rapidly grown to 128 percent of receipts as of the end of 2006.

Part of the explanation for the growing backlog of cases is the inability of the agency to increase the number of ALJs. From 1998 to 2004, annual hearing office receipts climbed from 458,000 to 576,000 while ALJs available to conduct hearings dropped from 1087 to 944. In other words, the number of new cases per ALJ increased by about 200 per year over that period. And the shortage of ALJs was aggravated by a shortage of support staff. So, what you had was rising workload and shrinking staff – an obvious recipe for creating backlogs.

What this has meant for the disabled individual seeking benefits is a much longer wait for the decision he or she is entitled to under the law. In 1998, when the Advisory Board wrote about the slowness of the process, the average claim in hearing offices had been there for an unacceptable 235 days, now the average age of a pending claim is 324 days – three months worse.

Again looking back to 1998, one of the striking statistics was the disconnect between the very high levels of accuracy reported at the initial determination level and the high rate at which initial denials were converted to allowances at the hearings level. Quality reviews of DDS determinations routinely found accuracy rates in the mid 90 percent range. But, about 62 percent of all hearing decisions turned DDS denials into hearing process allowances. Nine years later, quality review accuracy rates for the State agencies remain at the mid 90 percent level, but the hearing office decisional allowance rate has steadily increased, rising over the past 9 years by 10 percentage points to 72 percent at the end of 2006.

In most operations, the reaction to such discrepancies would be a ramping up of quality review efforts. The opposite seems to have happened here, at least at the hearings and appeals level. The review of hearings decisions in SSA has traditionally been much more limited than its review at other levels. A small sample of cases falling into certain error-prone categories were reviewed prior to effectuation and referred for further review to the Appeals Council. A second type of review, again a small sample of about 7,000 cases over a 2-year period but more representative of the general case load, was conducted on a peer review basis by volunteer Administrative Law Judges. It is my understanding that both of these review systems have been discontinued because of caseload pressures.

So we are left, even more so than 9 years ago, to rely on speculation as to why there is such inconsistency in outcomes. There are many theories. Some are benign. Many conditions are progressive and a claimant who was, in fact, ineligible at the initial decision may have become more disabled by the time of the hearing. Others are less benign. During the course of our work, the Board talks with a host of individuals involved in the adjudicative

process - claimants and their representatives, administrative law judges, DDS personnel, SSA employees and managers. We hear a lot of conflicting theories.

Some in the State agencies and some in the ALJ ranks maintain that the other group is misapplying the law. Some tell us that the pressures of the caseload lead to undue incentives to deny at the State agency level or to allow at the hearings level. A few years ago, the Board talked with an administrative law judge who told us that he tends to give claimants the benefit of the doubt. He went on to say that he knows other ALJs who tend to do the opposite. Then he said "And that's ok too." As a theory of adjudication, that is not ok in either direction.

Of course it is inevitable at any level, whether DDS or hearing office that the individual decision maker's background and attitudes may exert subtle influences that he or she may not even be aware of. But the frame of reference used in adjudication must be consistent. That is why you need quality review systems to assure that decisions are being made, with the greatest possible uniformity, in accordance with the standards established by the agency.

When you find substantial deviation from those standards, it is the agency's obligation to determine why that deviation exists and, insofar as possible, to cure it. The cure might be a refinement or clarification of the standards. The cure might be generalized training. The cure might be working with the adjudicators who are repeatedly making decisions that deviate from established guidelines and providing them with individual training.

This is not a new or novel concept. In 1980, Congress addressed this issue by enacting a statutory requirement for the agency to review hearing decisions. This provision of law has not been repealed. However, this law is rendered pointless if the reviews are not used to improve deficiencies.

Before the quality assurance program was suspended, agency reviewers found that the sampled cases were erroneous – either the wrong decision or insufficient evidence under agency standards –

in more than 4 out of 10 cases. Similar reviews applied to DDS level cases find erroneous decisions in less than 1 out of 10 cases.

In the peer review system I mentioned earlier volunteer ALJs examine hearing decisions under the looser “substantial evidence” standard. Under that review a case may not meet the evidentiary standards established by the agency but is not considered erroneous if there is at least more than the most minimal amount of evidence supporting the decision. Even under this much looser standard, the error rate at the hearings level is roughly double the error rate under the stricter standard applied at the DDS level.

When these numbers were mentioned at a recent Board meeting, one of our newer members asked a very salient question. Why shouldn't we expect a higher level of accuracy from those adjudicators who are more highly paid? The average personnel cost in ODAR is roughly 50 percent higher than that at the DDS level and ALJs are paid at levels generally comparable to senior Federal executives.

The results I have just mentioned are based on sample surveys. I have no reason to doubt that they were carefully designed and carried out. But sometimes, it is useful to do a reality check by looking at the full caseload to see if the big numbers are consistent with the findings of the sample survey.

I did a statistical analysis of the outcomes of hearings in fiscal year 2006 to see if the data told a story and they did. If you array administrative law judges by the number of cases they disposed of in 2006 and by the outcome of those cases, you see several things. First of all the range of cases handled and the range of allowance rates are both very wide. About a quarter of all judges disposed of fewer than 360 cases and 14 percent disposed of fewer than 240 cases. Half the ALJs disposed of between 30 and 50 cases a month during 2006 and average for all ALJs was between 400 and 500 cases per year. And the spread also extends on the upper side with about 10 percent of ALJs handling more than 720 cases in 2006. There are some ALJs who rendered decisions at incredible rates of 1000, 1800, and even 2500.

The average allowance rate of all cases disposed of in 2006 was about 60 percent and that is about the average for ALJs who handled 400 to 600 cases that year. Averages, however, hide the real questions about the decision making process behind them. Among judges who heard between 240 and 720 cases in 2006, the allowance rates varied from 3 percent to 99 percent. Among these judges who handled most of the caseload in 2006, 1.25 percent allowed less than 20 percent of the cases they ruled on in 2006 and 7 percent allowed more than 80 percent of their cases. I cannot believe that either the low or high allowance rates noted here are appropriate.

But judges who handle many more cases than the average tend to have significantly higher allowance rates, nearly 20 percentage points higher in the cases of those judges who dispose of more than 1000 cases per year. The raw statistics here cry out for more scrutiny regarding how cases are being handled across the organization.

I know that there are many anecdotal reasons advanced that purport to explain apparently anomalous numbers. Perhaps some of those arguments are going through your minds right now. But, this program is too important both to the taxpayer and to the affected individuals to dismiss statistical evidence with offhand theoretical arguments. There are administrative law judges who are deciding upwards of 1000 cases with allowance rates in the mid to high 90s. And there are administrative law judges who are deciding upwards of 1000 cases with allowance rates in the mid to low 30s. This is not a penny-ante poker game where we can shrug "Them's the breaks."

There is more in play here than decisional independence. If wrong decisions are being made then we are either depriving disabled individuals of vital income support and health insurance or we are improperly imposing on taxpayers a major cost that has been estimated to have a present value of about a quarter of a million dollars per case. And the numbers I see make it look very much like we are doing both to a completely unacceptable degree.

I know that much of the problem is attributable to factors beyond the control of individual administrative law judges. Clearly, inadequate resource levels have contributed greatly to the growing backlogs. They in turn create pressures and distortions that undermine administrative excellence.

I know that an unfortunate dispute between two other government agencies caused a freeze in hiring administrative law judges just at the time the agency was seeing rising caseloads that should have led to an expansion of capacity in the hearings offices. The Social Security Advisory Board has urged Congress to provide adequate resources, unfortunately with limited success. We have recently released an issue brief addressing the problems that SSA has faced in recruiting sufficient Administrative Law Judges and recommending changes to prevent a recurrence.

But, the reality is that caseloads are going to continue to grow and resources are going to continue to be scarce. This is not a situation that can be resolved by platitudes or a pep talk – although the latter may be useful from time to time. Fundamentally, however, it means that careful management will be needed. This involves better distribution of resources – there is nearly a 200 day discrepancy in average hearing processing time from the region with the lowest average to the region with the highest. And it involves the collection and use of data to pinpoint problem areas and to address them.

In a program with backlogs of 750,000 claims, the agency cannot afford to simply ignore the fact that some administrative law judges are producing at unacceptably low levels. It also needs to monitor high producing judges to make sure they are not sacrificing quality for speed. And it also needs to worry about judges who seem overly prone to allow claims or overly prone to deny them. While resource shortages create difficult tradeoffs, the agency must not simply give up on its responsibility for quality reviews and training.

There are, of course, special statutory safeguards that apply to the relationships between the agency and those of its employees who hold the position of administrative law judge. These provisions are

designed to assure that claimants are protected from improper efforts to influence decision making. They are not designed to allow administrative law judges a free pass on any oversight of their productivity, competence, and adherence to properly promulgated agency adjudication standards.

Having cited these hard numbers, let me also say that I fully believe that most Social Security Administrative law judges are conscientious, hard-working, and fully committed to producing hearing decisions that accurately apply the law and regulations to the particular facts of each case. And I suspect that that characterization is even more widely applicable to those of you here who have availed yourselves of the opportunity to participate in this training conference.

However, the fact that most administrative law judges are admirable employees does not mean that the agency should ignore its stewardship responsibilities here. It seems to me that ALJs who are carrying out their duties in a responsible manner should want to work with the management of the agency to develop reasonable standards and procedures that will fully protect decisional independence while identifying and seeking appropriate correction of situations where there is a failure to meet those standards. This can be through training, counseling, or, as necessary, referral to the Merit System Protection Board. Achieving a workable system of standards and procedures will not be simple.

But not trying to do so would be a dereliction of duty. Failure to establish standards and take appropriate actions to see that they are met is, quite simply, unfair. It is unfair at a very basic level to those who work in the program but have no benchmarks against which to measure how much is expected and whether they are doing well or whether they need to improve their performance. It is unfair to the taxpayers who have a right to know that their taxes are being used in accordance with federal law. And, it is especially unfair to claimants who, in a system without standards, have no basis for expecting efficient, timely, and accurate adjudication of their claims.

I know that you all have a very difficult job to do handling a high volume of complex claims with inadequate levels of support resources. I am confident that the large majority of those who work in this process do so with great dedication and care. The Social Security Advisory Board has urged more adequate resources and will continue to do so. But given the dual realities of limited resources and increasing caseloads, careful management to assure effective and efficient use of those resources is essential.