

ADMISSIBILITY OF EXPERT PSYCHOLOGICAL TESTIMONY
IN THE ERA OF DAUBERT: THE CASE OF HEDONIC DAMAGES

Gregory W. Joseph, Psy.D., Elliot L. Atkins, Ed.D.
and David K. Flaks, J.D., Psy.D.

The authors examine the issue of forensic psychological assessment of hedonic damages in light of the Supreme Court's decision in Daubert and beyond. Following a review of the rules of admissibility of psychological evidence, the authors address hedonic damages in this context. Case examples are used to illustrate observed weaknesses in both the scientific methods for assessing hedonic loss and the resulting testimony by forensic psychologists.

Increasingly, forensic psychologists are being called upon to evaluate and offer opinions about plaintiffs' claims of loss of pleasure in their lives as the result of an injury. In this area of hedonic damages assessment there is a growing debate among psychologists about the extent to which a subjective phenomenon, loss of pleasure in one's life, can be assessed with any accuracy and, further, quantified for the process of determining the actual monetary loss sustained by a plaintiff. The implications of having a scientific expert quantify such losses for a jury could be profound since awards for hedonic damages will presumably be based on this expert's numbers. In weighing the value of psychological evidence concerning hedonic damages, mental health professionals, as well as the courts, should be knowledgeable about the quality of the methods and procedures used in the formation of professional opinions. In that regard, the United States Supreme Court's 1993 decision, *Daubert v. Merrell Dow Pharmaceuticals* (1), is quite relevant in its re-examination of the criteria for admission of expert testimony in legal proceedings.

This article is designed to familiarize forensic psychologists with pertinent professional and legal issues regarding the presentation of

expert testimony in general and, through example, in the specific case of hedonic damage assessment. First, the *Daubert* decision will be reviewed as it pertains to the admissibility of expert testimony. Second, the relevance of the *Daubert* criteria to the presentation of expert opinions regarding hedonic damages assessment will be explored. Synthesizing these two areas, specific examples will be utilized to illustrate the authors' opposition to the introduction of psychological hedonic damage testimony in legal proceedings when the methodology behind it lacks a sound basis in the forensic psychological discipline.

HISTORY OF FORENSIC PSYCHOLOGICAL EVIDENCE IN THE COURTROOM

While the use of psychological evidence in legal proceedings is a relatively recent phenomenon, the courts are looking increasingly to the information and opinions offered by forensic psychologists. The use of such testimony is not limited to the courtroom or even to disputes between parties. Psychologists are also called upon to inform the courts via amicus briefs and other means. Moreover, the types of testimony offered by psychologists vary widely. In addition to scientific evidence, psychologists are presenting nonscientific testimony, as well as serving as fact witnesses for their clinical patients involved in civil matters, such as personal injury or child custody litigation. Before responding to a request to present psychological evidence, the nature of the type of evidence and the role of the psychologist should be explicitly delineated. The specialty guidelines for forensic psychologists, produced jointly by the American Psychology-Law Society and Division 41 of the American Psychological Association in 1991, are particularly useful in clarifying these issues.

While the *Daubert* decision does not specifically address the presentation of psychological evidence, nevertheless, it is relevant as it applies to the presentation of scientific evidence in general. A review by Goodman-Delahunty (2) revealed an increasing number of cases involving social and behavioral science testimony that cited *Daubert*

criteria. Prior to *Daubert*, courts relied primarily upon the standards of evidence set forth in *Frye v. United States* (3), a 1923 case in which evidence from a lie detector test was excluded because the court felt that the scientific principles upon which the test was based were not well recognized within the scientific community. The *Frye* decision established "general acceptance" among members of the scientific community as a prerequisite for admission of scientific evidence.

In 1975, Congress enacted the Federal Rules of Evidence (FRE) (4) which had been in existence for many years prior and specified standards for the admissibility of evidence. In addition to various common law standards, the Federal Rules of Evidence were subsequently adopted by a majority of the states, thereby replacing the standards set forth in *Frye*. Of particular relevance are Federal Rules of Evidence 401, 403, 701, 702, and 704.

FRE 702 allows for the presentation of testimony regarding scientific, technical or other specialized knowledge which is deemed helpful to the fact finder. It further states that the witness must qualify as an expert based upon his or her knowledge, skill, experience, training, or education. Under the *Frye* doctrine, only expert evidence which was outside the range of common knowledge and experience of the jury was considered admissible. However, this was regarded as incompatible with FRE 702's standard of assisting the trier. That is, the jury might still be edified and informed by the presentation and explanation of information which was, nevertheless, considered to be within its realm of knowledge and experience. Therefore, under the Federal Rules of Evidence this requirement was eliminated.

FRE 704 allows experts to offer statements about the ultimate issue in a case. However, this is frequently rejected in deference to common law rules in many jurisdictions which prohibit an expert witness from opining on the ultimate issue. Testimony which addresses the ultimate issue is often regarded as usurping the function of the trier of fact, and it is therefore disallowed. Somewhat peripheral for our

purposes, FRE 701 relates to percipient witnesses and allows for opinions or inferences based upon the witness's perceptions, rather than scientific theory or reasoning, even if the witness is not formally qualified as an expert. This rule is pertinent for treating therapists who can assist the court by offering their perceptions about a patient, even if they are not qualified as experts in the instant matter. By presenting the therapist as a nonscientific expert, counsel can cite this rule to circumvent efforts to exclude the therapist's testimony under the *Daubert* standards.

FRE 401 relates to the generalizability of a scientific principle to the evidence being presented in a specific case. That is, a direct connection must be shown to exist between the evidence presented in a particular case and the scientific principles purported to underlie that evidence. FRE 403 focuses upon the extent to which the scientific evidence proffered is informative rather than prejudicial to the judge and jury in deciding a matter.

THE DAUBERT DECISION

In 1993, the U.S. Supreme Court elucidated guidelines to be used in determining the admissibility of scientific evidence in *Daubert v. Merrell Dow Pharmaceuticals*. The case involved a suit brought against Merrell Dow Pharmaceuticals on behalf of Jason Daubert and Eric Schuller, both of whom were born with limb-reduction birth defects. The suit alleged that the birth defects were the direct result of their mothers' ingestion of Bendectin, a drug manufactured and sold by Merrell Dow. Both sides presented scientific evidence in support of their contentions regarding the role that Bendectin played in causing the birth defects. The trial judge rejected plaintiffs' expert evidence, opining that it failed to meet the general acceptance standard set forth in *Frye*. Arguing before the Supreme Court, plaintiffs maintained that the Federal Rules had superseded *Frye* as the standard for admissibility. The Court ruled that FRE 702 did address admissibility and it rejected the defendants' general acceptance argument:

The drafting history makes no mention of *Frye*, and a rigid "general acceptance" requirement would be at odds with the "liberal thrust" of the Federal Rules and their "general approach to relaxing the traditional barriers to opinion testimony" (1).

In so doing, the Court overturned *Frye* and incorporated aspects of the Federal Rules in establishing a new set of criteria for courts to determine the admissibility of evidence. Ironically, while its intention was to provide a more flexible standard of admissibility, the Court's opinion was actually more specific in its delineation of the criteria for admissibility. The intention was to bestow upon trial court judges a "gatekeeping" function in which they would become responsible for ensuring that scientific evidence is both relevant and reliable, as judged by the methods and procedures commonly employed by professionals in the relevant scientific discipline.

In its opinion, the Supreme Court specified three elements for determining the admissibility of evidence: reliability, relevance, and legal sufficiency. These were derived from the principles already in use under FRE 701-704 and 401 and 403.

Reliability

Although the *Daubert* opinion incorporates the concept of reliability already familiar to psychologists and scientists, the Court also espoused a broader definition in its discussion of the construct. The Court ruled that the testimony offered must be "scientific knowledge." It further defined "scientific" as being grounded in the "methods and procedures of science," and "knowledge" as "more than subjective belief or unsupported speculation." Regarding scientific validity, the Court stated that the principle should support what it purports to show. It further added that the focus should be on the principles and methodologies involved rather than the conclusions they yield.

To assist trial court judges with their new role of assessing the reliability of scientific evidence, the Supreme Court provided four guiding principles, which it regarded as nonexhaustive. First, is the theory or hypothesis testable or falsifiable? Interestingly, the Court only advocated that the theory be testable, not that it actually have been tested. Presumably, the idea is that the theory being proffered as expert scientific evidence must be amenable to the rigorous scrutiny of sound scientific inquiry.

Second, have the scientific findings been the subject of peer review and publication? While the Court appreciated that some scientific findings would be either too novel or of too narrow an interest to have been published, it regarded the review of scientific theories and findings by one's peers in the scientific community as an important safeguard, promoting the detection of weaknesses in methodology and procedure.

Third, is there a known or potential error rate associated with applications of the theory? Beyond the mere existence of a known error rate, this principle provides for inquiry about the statistical significance of a particular finding. Thus, while both parties to a dispute might agree about a particular method of measuring a psychological effect, based upon the error rate of the measurement, they might disagree about the statistical significance of the findings. Beyond admissibility, such an inquiry could affect what the jury might be told about the statistical significance of the study or how they should regard the findings based upon the range of error.

Fourth, is the technique or methodology being used generally accepted within the scientific community? While this guideline clearly relates to the issue of general acceptance, the Court explicitly stated in the *Daubert* decision that the Federal Rules of Evidence displaced the "general acceptance" standard for admissibility established in 1923 in *Frye v. United States*:

Widespread acceptance (of a scientific technique or theory) can be an important factor in ruling particular evidence admissible, and a known technique that has been able to attract only minimal support within the community may properly be viewed with skepticism (1).

Again, the Court's focus in this guideline is upon the methodology and/or scientific principle rather than the conclusions it generates.

Relevance

The second criterion, relevance, derives from FRE 401, and pertains to whether the testimony being proffered assists the trier of fact. The Court held that, in accordance with FRE 401, the scientific evidence must assist the trier of fact in understanding the evidence or determining a fact in issue. Moreover, there must be a clear link between the scientific findings presented and the case in question.

Legal Sufficiency

The third criterion set forth by the Court in *Daubert* is adopted from FRE 403 and concerns itself with the usefulness of the evidence in informing the trier of fact. That is, even though the scientific evidence is deemed by the trial court to be reliable and relevant, this third criterion asks whether the information proffered is more prejudicial than probative. The concern here is the exclusion of expert testimony which would mislead, prejudice or confuse the jury rather than assist them through the provision of useful information.

ADMISSIBILITY OF EVIDENCE FOLLOWING THE *DAUBERT* DECISION

Despite the Supreme Court's intention for its decision in *Daubert* to clarify and liberalize the criteria for inclusion of expert testimony, there has been considerable debate about whether the decision has made it easier or more difficult to have expert testimony admitted. As mentioned, the Supreme Court's decision in *Daubert* pertains only to

federal courts and is not binding upon courts at the state level. As a result, individual states have varied in their incorporation or rejection of the *Daubert* criteria. State supreme courts specifically rejecting *Daubert* in favor of retaining *Frye* are Arizona, Florida, Nebraska, and Washington. States which have adopted *Daubert* include Louisiana, Vermont, and West Virginia. Finally, still other states have endorsed the principles behind *Daubert* without explicitly accepting the language or specific criteria of the Supreme Court's decision. They include Arkansas, Delaware, Iowa, New Mexico, and Oregon.

Predictably, controversies have emerged about the interpretation and application of these standards. In his dissent from the *Daubert* decision, Chief Justice Rehnquist foresaw a particularly problematic issue regarding when to apply the four criteria for scientific evidence. He pointed out, for instance, that the decision was unclear about whether an expert seeking to testify on the basis of technical or other specialized knowledge should be held to the standards set forth in the "general observations" section which seemed to be limited only to "scientific knowledge." In general, forensic psychologists will be offering testimony which is considered scientific and is thus subject to the standards set forth in *Daubert*. However, the issue of "hard" versus "soft" science, particularly in the case of the behavioral and social sciences, has made it more challenging for forensic experts seeking to present psychological evidence in the courts.

In March, 1999, in *Kumho Tire Company, Ltd. et al. v. Carmichael, et al.* (5), the Supreme Court addressed this and other issues pertaining to the applicability of the criteria set forth in *Daubert*. The case involved expert testimony from Dennis Carlson regarding his analysis of the tire failure in an automobile accident which resulted in death and serious injuries to the plaintiffs. Defendants argued that the testimony should be excluded because Carlson's testimony was based upon technical information about the tires rather than scientific evidence. The Court found that in addition to scientific evidence, *Daubert*

also applied to testimony based on "technical" and "other specialized" knowledge.

The Court further opined that the *Daubert* criteria should be applied flexibly at the discretion of the judge. It stated that the criteria are neither necessary nor exclusive in determining whether specific testimony should be admitted:

The conclusion, in our view, is that we can neither rule out, nor rule in, for all cases and for all time the applicability of the factors mentioned in *Daubert*, nor can we now do so for subsets of cases categorized by category of expert or by kind of evidence. Too much depends upon the particular circumstances of the particular case at issue.

One implication of this decision for psychologists is the creation of greater latitude by the courts in recognizing psychological testimony as being within the purview of *Daubert*.

Forensic psychologists should generally be aware of the *Daubert* guidelines, and they must be prepared to demonstrate to the court the scientific and methodological underpinnings of their testimony. More importantly, forensic psychologists should be faithful to the standards of their profession. As mentioned, the Specialty Guidelines for Forensic Psychologists offer some direction to psychologists when functioning in the role of a scientific expert. Also, Marlowe (6) developed a hybrid decision framework for evaluating psychological evidence with regard not only to legal but to professional standards as well.

One outcome spawned by the *Daubert* decision has been an increase in the frequency of motions *in limine* according to FRE 104(a) in which opposing counsel seeks to determine the admissibility of expert testimony prior to trial or in seeking summary judgment. Because the proceedings from most of these pretrial hearings are not recorded, it has been difficult to ascertain the reasons behind their outcomes. Some decisions will be based on procedural grounds, such as the dis-

inction between scientific and nonscientific testimony. Others will stem from substantive considerations, such as the reliability and relevance of the scientific evidence itself. Still others will result from ill-informed or ill-prepared judges, lawyers, and/or experts.

While relatively few years have passed since the *Daubert* decision, there have been a growing number of challenges made to the admissibility of evidence in a variety of courts throughout the country. There has been a corresponding increase in the number of cases remanded on appeal back to the original court for reconsideration of admissibility under the *Daubert* guidelines. *United States v. Rincon* (7), a 1993 Supreme Court case involving eyewitness testimony, is illustrative of *Daubert*'s influence. In that case, the Supreme Court vacated the trial court's conviction, remanding the case for reconsideration of the admissibility of expert scientific testimony regarding eyewitness accounts. The trial court, followed by the Ninth Circuit, reviewed the case, upheld its original decision to exclude the testimony, and reinstated the conviction. Several reasons were specified by the Ninth Circuit. First, the defendant's expert failed to show that the testimony offered was related to an area of inquiry recognized as science. While vague statements were made about "the research," no specific studies were cited or presented to the court in support of their contention. Regarding a survey article of 63 experts presented on remand, the court opined that there was insufficient detail to determine if the study was scientifically valid. Second, following from FRE 403, the district court judge found that the testimony was more confusing and misleading than helpful in determining the verdict.

In a study which is currently ongoing, Bersoff et al. (8) have surveyed recent cases in the federal courts to determine the impact that *Daubert* is having. From their review, several areas pertaining to psychological testimony have been the subject of challenges to the admissibility of expert evidence. These include determination of sexual abuse, sexual offending, eyewitness identification, violation of civil

rights and discrimination, violence and victimization (e.g., domestic abuse), mental condition and state of mind defenses, and truth-telling techniques.

Preliminary findings from the survey reveal that in a majority of cases in which expert psychological testimony is challenged, the court has ultimately rejected the evidence, citing various aspects of *Daubert*. The one exception has been in the area of violence and victims, which includes testimony about various battering syndromes related to domestic violence and abuse. In the majority of these cases, expert psychological testimony was admitted. These initial findings raise interesting questions about why testimony pertaining to some areas is rejected while testimony pertaining to other areas has been accepted. Pending final results from this survey, it will be important to ascertain the reasons for these differences. Perhaps, there has been more focus, and thus a more substantial empirical base, on some areas than others. Alternatively, decisions about admissibility could have more to do with the knowledge and preparation of the experts and attorneys in demonstrating to the courts the validity of their evidence.

Before turning to the area of hedonic damages assessment, it will be helpful to briefly investigate how psychologists have responded to the new challenges brought by *Daubert*. Regarding expert psychological testimony in employment discrimination cases involving claims of pain and suffering, Goodman-Delahunty and Foote (9) offer useful guidelines to psychologists for offering expert opinions which satisfy relevant *Daubert* criteria. In essence, they contend that in forming an opinion, the forensic psychologist should focus on two issues: 1) demonstrating the proximate cause of the injuries and 2) documenting from multiple sources the nature and extent of the psychological injuries. In so doing, they correctly recommend that experts must address data which both supports and counters the claim of psychological damages.

Goodman-Delahunty and Foote (9) also emphasize the importance of utilizing a sound method of assessment which is generally recognized by the professional community as acceptable for gathering information and forming inferences about the claimant. Again, it is the method, rather than the conclusions derived from that method, which are at issue in the determination of admissibility under *Daubert*. The authors delineate a list of factors which are ideally addressed by forensic psychologists in their assessments of employment discrimination claims. All of these pertain to a determination of whether injuries exist, how bad they are, and what caused them. Thus, the examining expert would look at various sources of data (e.g., work and medical records, clinical interviews, and psychological test data) to ascertain whether the injuries did, in fact, result from discrimination on the job, and if so, to assess their extent and severity.

Other sources of stress which could contribute to a stress reaction and resulting diagnosable psychological condition must be evaluated as well. These could include the impact of daily stressors (excluding those relating to the alleged discrimination) both inside and outside the work setting. Other factors for consideration could include the impact of reprisals and retaliation for the discrimination action, stress due to unemployment if the claimant was fired or suspended, and pre-existing psychopathology. While it might be very difficult, if not impossible, to definitively tease out the causes for the claimant's measured distress, the forensic psychologist's testimony, based upon sound scientific methods of inquiry, can be of great value in assisting the court in making such determinations. While the factual findings and opinions offered by forensic psychologists in such cases might not change necessarily as a result of *Daubert*, experts will now be increasingly called upon to demonstrate the scientific underpinnings of their methods in order to have their testimony admitted into court proceedings.

Beyond the recommendations outlined by Goodman-Delahunty and Foote, we would also emphasize the importance of how the data

and findings from these evaluations are reported by forensic psychologists. First, the sources of data should be clearly delineated. Second, any limitations in the quality and/or quantity of information gathered should be specified with regard to how it affects the certainty of the conclusions being drawn in the report. Any contradictory findings which might temper the conclusions should be clearly reported. Finally, opinions offered by forensic psychologists should be stated as such, not as fact. If estimates of loss are to be reported, it should be obvious to all readers that these are estimates, with associated error rates reported where appropriate.

HEDONIC DAMAGES

Another area in which psychological testimony has been offered to courts concerns the assessment of hedonic damages. Purportedly different from damages awarded for pain and suffering, hedonic damages relate to the loss of enjoyment of life that an individual has sustained, and can reasonably be expected to suffer in the future, as the direct result of an injury. For many years, courts have allowed complainants to sue for hedonic damages. In 1992, the Supreme Court ruled in *Molozof v. United States* (10) that separate damages for loss of enjoyment of life were legitimate and could be recovered if allowed by state law. While a fair amount of attention in the legal context has been focused on the issue of hedonic damages, considerably less has occurred among social scientists in general and psychologists in particular.

Most cases involving hedonic damages testimony have relied upon the presentation of evidence by economists. Stanley Smith, an economist, first suggested the concept of hedonic damages in 1985, when he presented expert testimony in *Sherrod v. Berry* (11). While the case was ultimately reversed on other grounds, numerous subsequent cases have had Smith's and others' expert testimony on hedonic damages excluded. In particular, Smith's evidence has been rejected in the following cases: *Mercado v. Ahmed* (12); *Ayers v. Robinson* (13); *Fetzer*

v. Wood (14); *Longman v. Allstate Insurance Company* (15); *Anderson Courvillon v. Nebraska Department of Social Services* (16); and *Talle v. Nebraska Department of Social Services* (17).

In many of these cases involving testimony by economists¹, the notion that hedonic damages can be calculated using a mathematical formula was often rejected, as it was for the calculation of pain and suffering damages by the Supreme Court in the 1966 case of *Beagle v. Vasold* (18). In a recent rejection of Smith's testimony, the California Court of Appeal in *Loth v. Truck-A-Way Corporation, et al.* reiterated this point:

A plaintiff's loss of enjoyment of life is not "a subject that is sufficiently beyond common experience that the opinion of an expert would assist the trier of fact." No amount of expert testimony on the value of life could possibly help a jury decide that difficult question. A life is not a stock, car, home, or other such item bought and sold in some marketplace.

Smith's impersonal method of valuing life assumes that for the most part, all lives have the same basic value. That has democratic appeal, but Smith used no democratic processes in reaching that conclusion or selecting which benchmark figures to consider in setting the baseline figure. There is no statute Smith could have turned to for guidance. Our Legislature has not decreed that all injured plaintiffs of the same age and with the same degree of disability should recover the same hedonic damages; nor has it assigned set values in tort cases for the loss of an eye, ear, limb, or life (19).

¹ An in-depth analysis and critique of economists' testimony in this area has been carried out by Jerome Staller, Ph.D. See, for example, "Hedonic Damages: Junk Science Goes to Court," "What Price Life? Hedonic 'Experts' Claim to Know," "Comment on the Accuracy and Usefulness of Hedonic Loss Estimates" and "Placing a Value on the Enjoyment of Life" (20-23).

The opinion continues:

Our present system of requiring the jury to determine, without the benefit of a mathematical formula, the amount of a general damages award is not without its faults. But unless and until the Legislature devises a method for computing pain and suffering damages, a plaintiff may not supply, through expert testimony or otherwise, her own formula for computing such damages. Just as no judge may give the jury a standard for determining pain and suffering damages, no expert may supply a formula for computing the value of life and, by extrapolation, the value of the loss of enjoyment of life. That calculation, at present, must be left to the sound discretion of the jury.

Rather than examining the narrower question of hedonic damages assessment within a judicial context, psychologists have generally concerned themselves more broadly with issues such as the measurement of life satisfaction (e.g., among patients with cancer or terminal illness) or emotional distress (e.g., depression, posttraumatic stress, or adjustment disorders). In 1990, Berlá, et al. (24) attempted to define loss of pleasure of life with the development of the Lost Pleasure of Life Scale (LPLS). This reflected a trend by some psychologists toward a qualitative and quantitative assessment of loss of life pleasure which was then incorporated into the methods of valuation being developed by economists such as Smith. The weaknesses identified in the economic formulae were equally, if not more, problematic in these psychological approaches.

Berlá referred to a "change in quality of lifestyle," in contrast to earlier definitions which focused on an assessment of psychological pain and suffering. In a subsequent article, Andrews, Meyer, and Berlá stated that their definition of loss of life pleasure includes changes in life satisfaction, impairment and disability, changes in role function, and psychological elements. Their construct rests on two assumptions:

1) Pleasure in life is related to the sum of all factors (positive and negative) in a person's life, and 2) loss, impairment, or injury has a negative effect on a person's pleasure and potential for pleasure because of the restrictiveness and/or diminished control (25; p.100).

To date, the courts have had considerable experience in ruling on the admissibility of scientific evidence regarding the assessment of hedonic damages. As mentioned, much of this has been concerned with expert testimony from economists and less so from psychologists. Decisions pre-dating *Daubert* have included *Foster v. Trafalgar House Oil and Gas* (26); *Mercado v. Ahmed* (12); and *Fetzer v. Wood* (14). Those since *Daubert* include the following: *Wilt v. Buracker* (27); *Longman v. Allstate Insurance Company* (15); *Montalvo v. Makeku* (28); *Hein v. Merck and Co.* (29); *Ayers v. Robinson* (13); *Anderson v. Nebraska Department of Social Services* (16); *Talle v. Nebraska Department of Social Services* (17); and *Loth v. Truck-A-Way Corporation* (19).

Our experience in this area has included more than a dozen cases involving claims for loss of enjoyment of life. While the methods of assessment have varied widely among cases, what has been most troubling is the contention by these experts that they were able to not only assess an individual's loss of pleasure of life from the alleged injury but also to predict its effects on the individual's future enjoyment. Beyond that, several have attempted to quantify these losses by assigning a precise percentage of loss of pleasure which is then correlated to a corresponding precise dollar amount for recovery. It is our contention that while forensic psychologists can inform the court by offering a comprehensive assessment of those qualitative factors which weigh for and against an individual's claim of hedonic damages, the quantification of such loss into an exact percentage and corresponding dollar value is neither supported by empirical research nor the scientific community at large. Moreover, the presentation of such exact figures

is potentially misleading and prejudicial to judges and juries in its appearance of offering more certainty than actually exists.

In 1996, *McGuire v. City of Santa Fe* (30) relied upon the *Daubert* criteria to exclude expert psychological testimony regarding hedonic damages. The case involved a city employee, Frank McGuire, who filed an employment discrimination suit against the city based upon his claim that he was fired due to his age. In arguing the damages suffered by Mr. McGuire, attorneys relied upon expert testimony regarding hedonic damages which were assessed by a psychologist, Patricia Murphy, Ph.D., using the Lost Pleasure of Life Scale, and calculated into a dollar value by an economist. Citing the *Daubert* decision, the Court rejected the testimony on several grounds.

First, the Court held that there was no widely accepted standard for uniformly measuring the value of lost life pleasure. Second, it found that the sociological and economic basis of the concept of hedonic damages has been subject to widespread criticism. Third, in response to the economist's claim that the moderate validity and reliability of the figures offered were "greater than chance," the Court was unimpressed, commenting that the numbers produced were only more reliable than "those which might be drawn out of a hat." Finally, drawing upon FRE 403, the Court held that the testimony on hedonic damages was not helpful in assisting the trier of fact in understanding the evidence of the case.

Robert G. Meyer, a psychologist and co-author of the LPLS, responded by pointing out that the LPLS, which was designed exclusively for use in personal injury litigation, was improperly used here in an employment discrimination action. He also pointed out that the economist who testified in *McGuire* was ill-prepared to defend the reliability and validity of the LPLS. Moreover, the economist based his testimony upon the findings and report of Dr. Murphy who did not testify at the trial (31).

CASE EXAMPLES

Recently, the authors were consulted to address the validity of plaintiffs' claims of hedonic damages in mass personal injury litigation pending in both state and federal courts in a western state.

In one case, hedonic damages were assessed by plaintiffs' expert psychologist and vocational rehabilitation specialist using a method referred to as the Loss of Quality of Life Assessment (LQL). The authors were able to identify conceptual weaknesses and numerous shortcomings in the overall approach and methodology employed by the plaintiffs' expert evaluators. The LQL assessment methodology purported to evaluate each plaintiff's pre- and post-incident functioning, and the degree to which the alleged injury impacted their functioning via both qualitative and quantitative measurements of loss of quality of life.

The evaluators' stated methodology for qualitative measurement, or assessing the relative importance of each plaintiff's loss of quality of life in a variety of areas, was to compare each plaintiff's subjective report to observable, measurable data gathered from records, contact with fellow professionals involved with the case, and interviews. In practice, however, the evaluators were selective in gathering "objective" data, such as medical records, and little to no objective data were used in reaching their conclusions. Instead, the evaluators were highly subjective and drew largely from plaintiffs' self-reports, taken from questionnaires and inventories, to document each individual's purported loss of quality of life. For example, the evaluators used direct quotations from the plaintiffs regarding their pre- and post-incident functioning in several areas, including physical activity, academic functioning, social functioning, self-esteem, activities of daily living, career, and psychological functioning. For each of these areas an evaluation was then made regarding the degree of impairment and the relative importance of that area of functioning in the individual's life. From these an overall evaluation was offered. The following is a hy-

pothetical example, typical of the psychological functioning section contained in the evaluators' reports:

1) Pre Symptoms/Incidents

- "I was very happy in my life."
- "I didn't have a care in the world."
- "I had good self-esteem."
- "I had much support in my life."

2) Post Symptoms/Incidents

- "Now it's hard to get up each day."
- "If I start to feel things are getting better, I always take a turn for the worst."
- "I have lost my passion for life."
- "I can't shake problems off like I used to."

3) Evaluation

- Degree of Impairment: Marked
- Relative Importance: High

Overall Evaluation: Marked

On the basis of that information, the evaluators then made an assessment regarding the degree of impairment and the relative importance of that area of functioning in the individual's life, assigning a value of low, medium, or high.

In the final section, a summary of the evaluations of overall impairment from each area of functioning is presented graphically on a scale. Along a continuum, from none to extreme, the expert places an arrow in the spot representing the plaintiff's level of impairment for that particular area. The evaluators provided no explanation or rationale as to how any of these ratings were derived. Moreover, within each section on the continuum, the expert provides no rationale for why the arrow indicator is placed toward the high or low end of that section

(e.g., the high versus low end of the section delineated as "marked" impairment).

From a methodological standpoint, the problems and weaknesses in this approach are numerous. The categories of functioning might represent areas reasonably presumed to be affected by such an injury. However, no systematic study was undertaken to determine if the categories adequately cover the domain or, conversely, if they overlap. No effort was made to operationally define the construct of "quality of life" which was only vaguely defined by the evaluators. And using primarily only information obtained through self-report from plaintiffs in litigation is fraught with problems. Nor is there any indication that the evaluators sought and obtained corroborating evidence from external sources to support their subjective statements. Such statements are completely subjective and were not evaluated for accuracy. Moreover, no standardized tests were used to corroborate the plaintiffs' claims of emotional distress and loss of quality of life.

At the same time, the evaluators clearly included only that information which supported the plaintiffs' claims of damages even when data to the contrary existed. In each case the authors examined, data contradicting the plaintiffs' claims were identified among the various records and reports supposedly reviewed by the evaluators in forming their opinions. Many times the omitted information pertained to pre-existing conditions and experiences of the plaintiff which the evaluator neglected to include. Among the types of information excluded were facts such as family problems and deaths; criminal histories of the plaintiff or his or her family; histories of severe abuse; alcoholism or other addiction; and other life stressors which clearly impact one's quality of life.

Conversely, many assessments contained information which tended to contradict the claims made by the plaintiffs. For example, the spouse of a plaintiff might offer comments similar to the following about their loved one:

Since losing the weight and exercising every day, she is more mentally and physically fit than ever.

My husband has always felt a sense of worth from his work.

He always excels at it.

She has always had a positive spirit.

Her medical problems have never gotten in the way of her ability to care for the kids and be a terrific wife and mother.

Comments such as these were often omitted in the reports by plaintiffs' experts and seem not to have been incorporated into the evaluations made about the plaintiffs' loss of quality of life.

Regarding the quantitative assessments, the evaluators offered no rationale regarding how they arrived at their ratings of the degree of impairment, relative importance, and overall evaluation for each plaintiff in each of the categories. These all appear to be subjective judgments. Yet the evaluators used these judgments to provide a definitive number, a percentage, in fact, reflecting the amount of loss of quality of life sustained by each plaintiff. Not only are these numbers derived from subjective ratings based on insufficient and, perhaps, biased information, but they are used improperly to produce exact percentages of loss. The evaluators violated basic rules concerning levels of measurement when they went from ranges to exact amounts in their calculations. Rather than state that they are offering their estimates of plaintiffs' loss of functioning, the evaluators claim to be able to provide the exact percentage of loss which is then converted into a dollar amount extrapolated over time. There is no empirical support for any of these numbers.

The evaluators attempt to justify their approach by referring to the *Guides to the Evaluation of Permanent Impairment* (32) developed by the AMA, the American Psychiatric Association, and the Social Security Administration. This source recommends guidelines for the assessment of impairment in both physical and mental functioning. In

short, it recommends gathering data from individual providers, institutions, and nonmedical sources to assess the individual's functioning in a variety of areas, such as activities of daily living, social functioning, and stress tolerance. The *Guides* also call for the use of a degree of impairment rating scale which has five categories: none, mild, moderate, marked, and extreme. The authors of plaintiffs' analysis also make reference to the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV) (33) and the *International Classification of Diseases* (ICD) (34), although no specific reason was provided as to why these sources were relevant to their assessment method. Moreover, the manner in which they are used to justify the assessment approach is fraught with problems.

First, in reference to the DSM-IV and ICD, these manuals are used to identify mental and physical disorders, as defined empirically using specific and observable behavioral criteria. They do not purport to measure changes in one's quality of life. Moreover, the diagnostic categories are, for the most part, used to assess the presence and absence of a disorder, rather than the degree of its severity. For those conditions where severity is measured, such as in major depressive disorder, the categories range from mild to severe, again based on specific criteria, and are not converted into exact percentages.

Regarding the evaluators' reference to the *Guides*, they cite the third edition, which was subsequently replaced by the fourth edition in 1993. Here again, this manual refers to the assessment of impairment based upon a specified procedure of data collection and interpretation which contributes to an informed judgment offered by an evaluator about the presence and level of impairment. Even if the *Guides* were relevant to the assessment of loss of enjoyment of life, they expressly prohibit the use of percentages in determining mental disorders:

A more persuasive argument is that, unlike the situations with some organ systems, there are no precise measures of impairment in mental disorders. The use of percentages implies a

certainty that does not exist, and the percentages are likely to be used inflexibly by adjudicators, who then are less likely to take into account the many factors that influence mental and behavioral impairment. Also, because no data exist that show the reliability of the impairment percentages, it would be difficult for *Guides* users to defend their use in administrative hearings. After considering this difficult matter, the Committee on Disability and Rehabilitation of the American Psychiatric Association advised the *Guides*' contributors against the use of percentages in the chapter on mental and behavioral disorders in the fourth edition (32; p. 301).

In another case, the creators of the Loss of Pleasure of Life Scale (LPLS) were retained to evaluate and provide expert testimony about the hedonic damages sustained by the plaintiffs. The evaluators developed the LPLS for assessing loss of life pleasure in the context of personal injury litigation, and describe their scale in two articles (24, 25). The first article proposed a conceptual framework for defining the concept of hedonic damages, measuring hedonic loss, and translating those findings into dollar values based upon economic theories in an interdisciplinary approach. The second article purports to offer reliability and validity data for the LPLS.

In defining the domain of hedonic damages, Berlá and his colleagues (24) utilized agreement among professionals to develop a list of categories pertaining to an individual's routine functioning. These included practical, emotional/psychological, social and occupational functioning. Behavioral descriptors were developed for each area of the domain, though they were not exhaustive for each level of severity. The reliability study involved rating the loss of pleasure in a series of 15 prepared vignettes using the LPLS. Interclass-correlations of .65 were obtained across all groups (psychology graduate students, clinic staff and psychologists) for short-term ratings and .70 for long-term ratings. The validity study involved having 27 undergraduate psychol-

ogy volunteers independently rate 15 one-page, prepared vignettes and award a sum of money for compensation of loss of pleasure in each case. Spearman rank order correlations of .86 to .88 were reported.

Like the previously described LQL assessment methodology, the LPLS methodology presents serious shortcomings which we contend make it ineffective for quantifying hedonic damages. First, Berlá and his colleagues offer no standardized method for gathering and combining data for consideration in making ratings with the LPLS. In their study, subjects were presented with prepared, one-page case vignettes. This differs vastly from the voluminous information through which a forensic psychologist typically must sort in arriving at opinions about a client. Moreover, there are no guidelines offered for how the various types of information should be weighted and combined. In the validity study, hypothetical jurors were again given prepared, one-page vignettes and did not engage in a process of deliberation as would occur in an actual case. Berlá himself is aware of these weaknesses, having acknowledged them in his article and in deposition testimony.

Similar to the evaluators using the LQL methodology in the previously described case, Berlá and his associates, in practice, selectively utilized available information and neglected to mention numerous facts, unrelated to the alleged injury, which were clearly relevant to the plaintiff's levels of life satisfaction. Typical of the sorts of pertinent information often not reported were facts such as numerous marriages; abusive relationships; convictions for serious crimes, such as manslaughter; psychiatric hospitalizations; inpatient drug and alcohol rehabilitation; and other, unrelated, serious medical problems and injuries. Each of the test cases reviewed in this action contained such excluded information which was clearly relevant in the assessment of the plaintiff's level of enjoyment of life. Again, the LPLS's lack of a standardized method of gathering and combining data renders it highly susceptible to such selectivity and omission biases.

Perhaps more important is the absence of any empirical evidence to show that the ratings produced using the LPLS bear any resemblance to actual loss of life pleasure sustained by a real plaintiff. It seems that such concurrent and predictive validity must be established before the scale can be used in nonempirical contexts such as personal injury litigation. Moreover, given the heightened potential for bias in the self-report of litigants, validation of the LPLS should ideally include comparisons of loss of life pleasure with nonplaintiffs and non-injured individuals. Also, the LPLS lacks any convergent or discriminant validity linking it to similar, established measures of psychological functioning and distress.

In calculating dollar values corresponding to hedonic damages sustained by litigants, the LPLS also exhibits flaws similar to those found in the LQL. Berlá and associates claimed that they were offering estimates of plaintiffs' loss of life pleasure, but, nevertheless, went from providing ranges (e.g., Mild, Moderate, Severe) to medians and, ultimately, to percentages. The ranges used were as follows: None (0%), Minimal (1-17%), Mild (17-33%), Moderate (33-50%), Severe (50-67%), Extreme (67-83%), and Catastrophic (83-100%). While their initial ratings were expressed as ranges of severity, they reasoned that they could then use the median value of the range to depict the amount of loss. This erroneously replaced an ordinal level of measurement with a more precise ratio level. These percentages, rather than the ranges, were then reported with no confidence intervals based on error rates, and without qualification that the percentages were merely estimates. Moreover, the percentages corresponding to the ranges overlapped, so that, for instance, a rating of 17% corresponded to both the Minimal and Mild ranges. Also the ranges differed in size, some encompassing 16 percentage points, others encompassing 17 percentage points. Even the descriptors for the ranges were potentially misleading in that half of the total continuum (i.e., all ratings of 50% or more) was considered to be Severe or worse.

In both of the above-described cases, as well as in several other cases for which the authors have been consulted to assess the validity of hedonic damage assessments performed by mental health professionals and others, we have consistently encountered the use of methods and procedures which lack any empirical basis. With the exception of the LPLS methodology, which at least has one empirical study as support, all of the other methods of assessment of hedonic damages have relied upon checklists and scales that lack any indicia of reliability and validity. Yet even the lone study regarding the LPLS methodology fails to demonstrate reliability and validity for use of the LPLS in real world cases. Moreover, while the findings from these assessment approaches, frequently erroneously converted into exact percentages of hedonic loss, have clearly lacked empirical support, their use in economic formulae for determining corresponding dollar values is even further afield from the rigors of scientific validation. There simply is no theoretically sound and empirically tested rationale to support these claims of quantification of hedonic loss. Scales, such as the LPLS, could be used by forensic psychologists as a tool for gathering and organizing data about an individual from an evaluation, but they have no other utility beyond this. To offer numerical estimates, let alone exact figures, about an individual's loss of pleasure of life clearly exceeds the proper use of such scales.

To summarize, these scales and assessment methods lack content validity in that they have not shown the degree to which the constructs encompass the domain of what most would consider relevant for hedonic damages. Moreover, they lack studies that operationalize and then validate what is meant by the construct of loss of life pleasure. Finally, to date no scales or methods of assessment of hedonic loss have demonstrated any criterion validity. There is no evidence of whether and how the findings from these scales and approaches resemble or differ from control and comparison groups such as noninjured plaintiffs and injured nonplaintiffs. Further, there are no pro-

spective studies to support the extrapolation of hedonic loss assessments into an individual's future; nor are there any corresponding error rates when such predictions are offered by plaintiffs' experts.

The reasons for such a dearth of empirical support are unclear. However, we suspect that the proposed utility of these scales and assessment methods in the context of personal injury litigation has primarily driven their development rather than scientific curiosity or clinical necessity. This raises yet another concern about the use of these scales and the concept of quantification of hedonic damages itself. Even if the scales could be shown to demonstrate reliability and validity, the nature of information considered in each forensic evaluation varies greatly. Moreover, the factors deemed highly relevant by one evaluator will likely differ, perhaps greatly, from those of another. One need only attend a trial involving a "battle of the experts" to see how mental health experts can vary widely in their clinical perceptions and opinions. Without a standardized means of gathering and weighing the data of a particular case, one can never be sure whether the results derived from such an assessment are truly comparable from one expert to the next. Fortunately, the adversarial nature of our judicial system, particularly with the assistance provided to the courts through the *Daubert* guidelines, provides a safeguard in allowing scrutiny of experts' findings and requiring that experts cite the bases for their perceptions and opinions.

Because results from hedonic damage assessments are being used to claim monetary damages in actual cases, the foundation upon which those findings are based must be theoretically sound and empirically tested. To date we know of no such support for the quantification of hedonic damages. Moreover, as we have shown through case examples, without a standardized method of assessment, there exists a significant potential for bias in the results when plaintiffs' self-reports are unduly relied upon, often to the exclusion of contradictory information. When exact percentages and corresponding dollar values are then