IDEA: The Journal of Law and Technology
1992

SCIENCE AND SCIENTIFIC EVIDENCE

Steven J. Grossman [n.a]

Abstract

This study contemplates what the scientific community judges as being good science, compares that to existing legal standards, and suggests a change in the perspective concerning the admissability of scientific evidence. Since the drafters of the Federal Rules of Evidence could not reach agreement on the proper role of the Frye standard for the admissability of scientific evidence, common law development has centered on some sort of consensual basis to decide whether or not a given piece of scientific evidence should be admitted. By considering what the scientific community judges as being good or bad science, science reveals itself as a method of prediction through the use of empirically verified models. Such models do not become verified as a consequence of social powers (consensus). Models become verified when their predictions are subjected to the intensity of repeated and rigidly controlled experiments. Therefore, courts must begin to subordinate the notion that consensus has an important role in establishing scientific validity. Reliability of prediction and its associated relevance should be central to admissability issues.

*180 "We need to think long and hard about the future of society as technologically orientated and as law-soaked as ours when our scientists and lawyers cannot even talk to each other"

The above comment of the Honorable Howard T. Markey, former Chief Judge, United States Court of Appeals for the Federal Circuit [n.1], describes an issue that appears to get to the heart of the current debate concerning the utility, admissability and proper role of scientific evidence in modern day litigation. In practice, the law of evidence provides systems of filters sensitive to the balance of aiding the factfinding process versus many of the well-known countervailing considerations (undue consumption of time, danger of unfair prejudice, confusion of the issues, misleading the jury, and unfairly or harmfully surprising a party who has not had a reasonable opportunity to anticipate the evidence submitted). [n.2] In the context of evidence that is considered scientific, courts then have an obligation to exclude bad science, that is information based on scientific techniques which are simply wrong. After all, bad science would necessarily raise some of the mentioned countervailing considerations to a point that outweigh any advantage in aiding the factfinding process. But what is bad science? Who defines science? When does evidence become scientific and not plainly demonstrable? By ignoring the challenge of Chief Judge Markey answers to the above have led to some unusual and inconsistent policies concerning the role of scientific evidence. [n.3] This study begins to contemplate what the scientific community judges as being good science, compares that to the existing legal *181 standards, and suggests a change in the perspective concerning the issue of admissability of scientific evidence. [n.4]

The Frye Rule

In 1923 the District of Columbia Circuit Court developed what is widely known as the Frye rule for determining the credibility or reliability of novel scientific techniques when ruling on its admissability. [n.5] The notion at that time was that scientific evidence needed to be treated differently than other types of evidence for several reasons, most importantly of which was the aura of mystic infallibility such evidence was believed to hold with the average juror. As a result, the principle danger of scientific evidence was considered to be its potential to mislead the jury. [n.6] The widely-quoted language delineated in the Frye test is as follows:

Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognized, and while courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, the thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs. [n.7]

The Frye test is usually construed as necessitating a survey and categorization of the subjective views of a number of scientists, assuring thereby a reserve of experts available to testify. [n.8] Consequently, scientific evidence is then admitted when viewed as generally accepted by the scientific community in which it belongs. Of course the irony here is that the court is judging the admissability but almost never will have sufficient personal expertise to evaluate the validity of any scientific conclusion. [n.9] This has led to numerous discussions concerning the validity of the Frye rule.

*182 Criticism of Frye

As noted by recent reviews, it is probably a fair comment to say that over the years the Frye standard has been warmly embraced, emphatically rejected, occasionally ignored, and openly modified. [n.10] While there exists an extraordinary amount of law review comment, [n.11] at a minimum it is fair to say that where proffered evidence arises from a novel form of scientific expertise, the courts have taken one of three different approaches to the question of admissability. [n.12]

First, there are those courts that have adhered to the Frye standard, requiring the "general acceptance" standard as dispositive to the issue of admissability. [n.13] Studies completed in 1983 suggested that the Frye rule continued to have its supporters and could be viewed as the majority rule. [n.14]

The courts that have abandoned Frye have to a large degree focused on the possibility that under Frye, recent advances of science that have a high degree of precision and accuracy, [n.15] yet are so novel that there are only a few adherents, would be unnecessarily excluded from the fact finding process. More candidly, these courts have argued that a determination of reliability cannot rest solely on a process of "counting scientific noses". Furthermore, unanimity of opinions in the scientific community, on virtually any scientific question, is extremely rare. [n.16] Courts following such lines of argument have therefore considered scientific evidence as admissable as long as there was recognition by a specialty within a general field, that is, "substantial acceptance" as opposed to "general acceptance" within the relevant scientific community. [n.17] The ultimate view emerging was that Frye frustrated, rather than enhanced, the search for truth. [n.18]

*183 The third group, and to some extent the most liberal view, are those simply favoring a relevancy standard. [n.19] This rationale has focused on Federal Rules of Evidence 401 and 402 (relevancy), 403 (relevancy balanced against unfair prejudice) and 702 (expert testimony). Under such a view, the Federal Rules are construed as a implicit rejection of Frye and its general acceptance standard. [n.20] In effect, if a piece of scientific evidence provides any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable, and there is no danger of substantial prejudice to the defendant, and the proffered scientific evidence via some expert will assist the trier to understand a fact in issue, the scientific evidence is admitted.

Judicial Attempts To Solve The Frye Debate

In the face of all of the above viewpoints, at least one court has tried to distill a somewhat modified version of the Frye standard while keeping in mind Fed.R.Evid. 702 [n.21] and the general spirit of the Federal Rules of Evidence. In United States v. Downing, a detailed analysis concluded that the Frye test suffered from serious flaws and reflects a conservative approach to the admissability of scientific evidence that was at odds with the spirit, if not the precise language, of the Federal Rules of Evidence. [n.22] A particular degree of acceptance of a scientific technique within a scientific community was viewed as neither a necessary or sufficient condition for admissability. [n.23]

*184 Consequently, the admissability of scientific evidence was found to turn on a consideration of (1) the soundness and reliability of the process or technique used in generating the evidence, (2) the possibility that admitting the evidence would overwhelm, confuse, or mislead the jury, and (3) the proffered connection between the scientific research or test result to be presented and the particular factual disputes in the case. [n.24]

The determination of reliability was held to incorporate general scientific acceptance, as well as the existence of specialized literature dealing with the proffered scientific technique, the qualifications and professional statute of the expert witness, the non-judicial uses to which the scientific technique was put, the frequency with which a technique leads to erroneous results (i.e., a very low rate of error strongly indicating a high degree of reliability), the type of error generated by a technique, and finally judicial notice of expert testimony that has been offered in earlier cases to support or dispute the merits of a particular scientific procedure. [n.25] The court concluded that other factors could be added to the list. [n.26]

Assuming that some degree of reliability is established, such must be balanced against the possibility of jury confusion. The chief concern here of the Downing court was that certain evidence, although bearing substantial indicia of reliability, could confuse rather than assist the jury. The court referred to those techniques that "assumed a posture of mythic infallibility" and those situations where the jury would be asked to accept the expert's assertions as to the accuracy of their conclusions. [n.27]

*185 The final consideration under the Downing test is whether the expert testimony proffered is sufficiently tied to the facts of the case so that it would aid the jury in determining a factual dispute. [n.28] Consequently, the party seeking admission of scientific evidence would satisfy this last consideration through an on-the-record offer of proof. [n.29]

Science and The Scientific Method--Creating Models

Science may be defined as an objective, logical and systematic method of analysis of phenomena, devised to permit the accumulation of reliable knowledge. [n.30] The immediate purpose of scientific thought is to make correct projection of events in nature. The essential notion behind the various phrases is the idea of prediction. A scientific fisherman, for example, is one who can predict successfully when and where the fish will bite. Every time the word science or scientific appears, the idea of prediction is explicitly or implicitly involved. [n.31]

Consequently, the whole effort of scientific study is directed toward one single end: verification of prediction. [n.32] This is generally achieved by postulating a model whose main purpose is to make such predictions. If the predictions are correct, the model is successful and is said to be validated. The meaning of validation is that the model has previously made correct predictions. Note that the word validated is preferred over the notion of proved, since the latter imports too much certainty and permanence to be used in describing models. Furthermore, the notion *186 of true and false is inappropriate since such statements apply only where empirical verification is possible. [n.33] A model is not said to be true or false; it is said to be validated or invalidated.

For example, a mathematical model refers to equations or other relationships that provide the quantitative predictions of the model. Consider a couple of mathematical models from the world of physics, namely, momentum is equal to mass multiplied by velocity and force is equal to mass multiplied by acceleration. By checking these models against the actual measurements or observations of events in nature, they have been well validated. In a given automobile accident, a scientist would be prone to employ these models to make some prediction as to how fast a particular vehicle was travelling at the moment of impact in the absence of direct evidence of actual speed. [n.34]

It is important to recognize at this point that the any characterization of scientific evidence must be consistent with the notion that scientific thought contemplates the postulation of a conceptual model of nature from which the observable behavior of nature may be predicted accurately. New observations and new measurements (that were not predictable) result in an adjustment of the model so that it can again be verified.

From this perspective it is useful to reflect on our understanding of classical scientific testimony. For example, the results of scientific tests of alcoholic content of a person's blood have generally been held admissable on the question of intoxication. [n.35] The operative model is that blood alcohol content will predict an individual's level of intoxication. The use of fingerprints for the purpose of identification is based on our understanding that the skin ridges which make up the fingerprint pattern remain unchanged from birth and until the skin decomposes after death. [n.36] The operative model is that every individual has a different pattern so that a single fingerprint impression can predict an individual's presence at some scene. The use of radar for speed detection has come so far as some state statutes provide for judicial notice of the scientific principles underlying its employment. [n.37] The operative *187 model, of course, is that a radio wave which strikes a moving object changes frequency proportionately, and allows for prediction as to the speed of the object. Finally, scientific evidence is often used to identify blood or bloodstains by chemical analysis or to distinguish between human and animal blood. [n.38] The operative model is that the human population can be classified into four major blood groups (O, A, B, and AB) along with the approximate percentage of person falling within each group. [n.39] Consequently, the model of blood typing allows for prediction but only to the limit provided by our current modelling of the percentage of persons falling within each group. Pursuant to new observations and new capabilities of measurements, new models of blood classification are being proposed that may provide for finer discriminations that distinguish the blood of one person from that of another. [n.40]

How Models Are Validated by Science

There are only two basic types of verification accepted in legitimate science. [n.41] The optimal form is prediction beyond pure chance. This is demonstrated by subjecting the model to empirical tests in some pilot study. In a typical case, a small sample is selected for testing which serves a dual purpose. First, the predictive value of the model is empirically established. Additionally, this allows for further modifications of the model in order to increase reliability, hence validity.

Quite simply, an empirical test of our model that predicts an individual's level of intoxication from blood testing is established by having a group of individuals ingest varying amounts of alcohol along with an analysis of their blood chemistry and behavioral patterns. An empirical test of our model of radio waves changing frequency upon striking a moving object would contemplate sending objects at different speeds through a radar field and analyzing the frequency shifts as a function of object speed.

An empirical test of our model that every individual has a different fingerprint pattern would be established by a showing that for a given number of individuals there is an equal number of patterns. An empirical test of our model of blood grouping would consider a sample of individuals and a determination of the percentages of individuals with the various blood groups.

*188 The second type of verification, which has been characterized at a lower logical and functional level than the above, relies on consensus. As employed in science, consensus has been described as one of three types: (a) agreement amount the established and accepted leaders in a given field (authority figures), (b) agreement among groups of authorities as defined by their followers within a select group ("schools of thought"), or (c) agreement by virtue of isolated declarations of a single person. But as noted, the idea of consensus is substantially inferior to verification by prediction and exists under the notion that "something is better than nothing". To some degree, verification by consensus is a system of circular proof. A model is proposed and is validated owing to the acceptance by a number of individuals who simply believe in its ability to make predictions. On the other hand this method of verification is not wholly unreliable. Consider that Einstein was confronted with an organized campaign of ridicule until his theories of relativity were empirically verified.

Recognized Problems in The Two Basic Types of Verification

A. Empirical Testing

From the above we have established that the key to empirical testing, that which strictly sets the bounds within which the validity of scientific modelling can be achieved, is objective measurement. Unfortunately, measurement can be objective only under rather stringent conditions.

First, before any measurement can be made, it is essential to have a standard unit for comparison. Next, the act of observation should not in itself alter the state of the object being observed. [n.42] It is also important to make sure what one is measuring. [n.43] These latter considerations support *189 the notion of accuracy. Furthermore, if objective validity is to be ensured, the measurement must be reproducible, that is precise.

This latter idea of reproducibility and precision was touched on earlier, and a distinction was made between this concept and the idea of accuracy. It is worth repeating that accuracy is reserved for the implications that are associated with some precise observation. Precision, on the other hand, can be thought of as the initial quality control measure that we apply to an act of observation. It only goes so far as suggesting that the observations have been consistently the same under the selected technique of testing. Consequently, empirical testing and validation of any model requires that both precision and accuracy are present. For example, consider the model of radio waves changing frequency upon striking a moving object. In an effort to empirically test this model an automobile is driven through a radar field and the speed is recorded. It turns out that a steady wind is blowing in the same direction of our vehicle, and the radar, unknown to the observer, and for the purpose of this hypothetical, responds to this wind current. Furthermore, the observer inadvertently moves the radar gun in a similar manner every time an observation is conducted. As the vehicle passes this radar field, at some given exertion on the gas pedal, the same or precise measure of speed is recorded. Although the act of observation is precise, the act of observation has altered the state of the object being observed and it is not certain what the radar is measuring (wind or object). Under these conditions, the model is not verified in the explicit sense, i.e., its ability to make accurate predictions concerning the speed of a particular moving object.

One can conclude from the above that in testing a scientific model. there must be great patience and an accumulation of vast amounts of observational data. [n.44] A single observation of some unique event is scientifically valueless. It is impossible to be sure that a single measurement is unaffected by some technical error or chance variable. Furthermore, observations are subject to bias if the scientist has any preconceived ideas as to what the result should be. [n.45] Objective measurement depends both on the object and the observer, and is not a property of the object alone. In sum, the criterion of the scientific status of a model is its testability. [n.46]

*190 B. Consensus Testing--Scientific Criterion of Consensibility

Supreme Court Justice Stewart wrote that perhaps he could not intelligently define hard-core pornography for criminal prosecution purposes, "But I know it when I see it". [n.47] It has been suggested that much the same is true of the scientific "consensus". [n.48] Quite simply, there are some important points of distinction between the idea of consensus as perceived and understood by scientists and disregarded by the courts.

As noted earlier, the idea of consensus has been legally defined as some form of a "general acceptance" standard. In science, it was noted that consensus was described as one of three types starting with agreement among established leaders, followed by agreement within select groups and finally agreement by virtue of isolated declarations of single persons.

At first glance, the comparison between legal and science type definitions may appear to be similar. Upon closer examination, however, the idea of scientific agreement takes on more complex meaning. First, when scientists say that they agree with a given model, it is more appropriate to say that they understand a given model and it is regarded as a special case of a more general model. [n.49] The inclusion of the successful parts of old models into some new and more general model, is typical of the history of science.

Furthermore, if agreement in science exists, with recognition of the ability of some new model to make correct projection of events in nature, it is a consequence of some amount of empirical testing. When consensus is present in science it is not present in place of empirical testing. Rather, consensus often represents an indicia of acceptance concerning the methods of empirical testing in relationship to some newly proposed model.

*191 For example, consider that some new model for blood typing is proposed which goes beyond blood grouping and allows for individual determination of blood source. In other words, according to the model, blood is as individual a trait as fingerprints. If this model includes the successful parts of the old models of blood typing and is offered with full empirical testing, there may be complete agreement about this model, full consensus owing to full empirical testing.

Alternatively, if some amount of empirical testing is demonstrated, there may be some consensus as to the validity of the model. Those scientists that take issue at this point could be divided into two groups. Those that still reject the underlying existence of such a model of blood typing, and those that take issue with the offered methods of empirical verification.

This latter group may not necessarily reject the possibility of the model, what is rejected is the method of verification. Arguably, there is a consensus here from the perspective that at least the possibility of the model has been agreed upon. Furthermore, this group may actually embrace parts of the model if parts of the verification methods are accepted. Here then there is a consensus as to the validity of a portion of the original proposed model. Alternatively, they may propose a modification of the model that is considered a more accurate explanation of the offered empirical testing. At this point we have a consensus for a new model as a consequence of empirically testing what is now a rejected model.

The former group can be considered our strongest skeptics. In addition to fully rejecting the possibility of the model, they have rejected the methods of empirical testing offered to validate the model. As far as this group is concerned, there is a consensus that no new model exists.

Finally, if no amount of empirical testing is offered we have the possibility that the model will be accepted subject to future verification, or, the model stands rejected with the idea that future verification is not possible, is not needed, or would be a waste of time. One could argue at this point that there is a consensus as to the possibility of a model with a consensus that empirical testing will eventually confirm the model. Alternatively, one could suggest that there is a consensus as to the possibility of the model but a consensus that the model will not be capable of empirical verification. Finally, one could claim that there is a consensus that the model is not possible and verification is impossible.

If consensus emerges without empirical testing, it is a warning sign. When consensus through rejection or acceptance of empirical testing is not the main objective, and there is a lack of mutual criticism and tacit cooperation towards a limited common basis of acceptable models, *192 it is probably a non-scientific body of information. A sure symptom of non-science is personal abuse and intolerance of the views of one scholar by another. [n.50]

Other lines of demarcation between science and non-science have been suggested. A model or discipline which purports to be scientific, has been termed pseudoscientific if and only if:

(1) It has been less progressive than alternative [models] over a long period of time, and faces many unsolved problems; but (2) the community of practitioners make little attempt to develop the [models] towards solutions of the problems, shows no concern for attempts to evaluate the [models] in relation to others, and is selective in considering confirmations and disconfirmations. [n.51]

For obvious reasons, it is somewhat risky to suggest that various disciplines belong on the non-science side of the above demarcation. Some commentators have suggested that such disciplines as literature, philosophy, sociology, social psychology, economics, history, astrology and theology appear to fit into the above non-science category. [n.52]

This is not to suggest that scientific knowledge is unachievable in some of the above disciplines (sociology, social psychology, economics). Rather, if such social scientists are truly scientific and not in too much of a hurry to assert final conclusions, verifiable models of social phenomena may become available. [n.53] In other words, there are some small emerging groups of social scientists who employ scientific methods (empirical verification) to generate predictive models, but within the broad field, for example, of economics, it is not a common practice. [n.54]

Legal Misunderstanding of Scientific Modeling, Verification and Consensibility

A. The Breathalyzer

Of all the tests to measure alcohol intoxication, the simplest and most socially acceptable one is the test of a specimen of breath. Expelled "deep *193 lung" or "alveolar" air has been suggested to accurately reflect that which a blood test might reveal. [n.55] Such tests have been characterized as "evidentiary tests" because the results have been considered sufficiently reliable to be admitted at a civil or criminal trial on the disputed issue of blood alcohol content. [n.56]

The operative model in the determination of intoxication from breath analysis employs the use of Henry's Law. At the risk of some oversimplification, this law predicts that if a volatile substance such as alcohol is dissolved in some solvent (blood), the concentration of the alcohol in the vapor above the liquid (i.e., your breath) will be proportional to its concentration in the blood. As part of this operative model, the concentration of alcohol in the blood has been postulated to be 2100 times that in the breath. [n.57]

Consequently, upon application of a breath analysis, our proposed model predicts that the alcohol measured in the breath results from ingestion and can be multiplied by a factor of 2100 to provide an accurate measure of the alcohol in the blood. A second operative model predicts states of inebriation from the predicted blood alcohol concentration. [n.58]

Since most states still hold that breath tests are sufficiently reliable to be admitted into evidence, it is worth reflecting on how this level of judicial acceptance compares with our new understanding of modeling, verification and scientific consensibility.

As noted earlier, it is critical to the validation of any model to understand the distinction between precision and accuracy in relationship to empirical testing. Putting aside other possible sources of error, if the breathalyzer machine is functioning properly, it provides a precise measure of alcohol vapor in the breath. That is, upon repeated testing, the breathalyzer will likely yield consistent observations of alcohol in the breath. One could also suggest that the breathalyzer result is *194 accurate in the sense that the implications from these precise measures of alcohol in the breath allow us to predict that the individual has ingested alcohol. This much of the proposed model appears fully validated.

But note that the breathalyzer is not accurate in the sense that the precise measure of alcohol in the breath allows for prediction of a blood alcohol content. The proposed model postulates a blood alcohol: breath alcohol ratio of 2100, but empirical verification of such a ratio for all human beings has not been established. [n.59] Consequently, in a recent Nebraska case, State v. Burling, the Nebraska Supreme Court ruled that the test results from some breathalyzer had to be reduced to 52.38 percent (of the reported value) to accommodate any possible variations in the blood alcohol:breath alcohol ratio. [n.60]

As to the issue of consensibility, it appears that what existed was something far less than a consensus that the "1:2100" model itself was verified. In March of 1952, under the auspices of the National Safety Council's Committee on Tests for Intoxication, the leading researchers on the subject signed a joint statement which suggested that "Available information indicated that [the] alveolar air-blood ratio is approximately 1:2100". [n.61] As noted by Prof. Erwin, such a statement on its face was ambiguous since it employed the word "approximate". [n.62] Nevertheless, it is fair to conclude that in 1952 there may have been a limited amount of empirical testing, with only a consensus as to the possibility of the "1:2100" model. Twenty years later, in a statement signed by some of the original members of the 1952 declaration, it was noted that "Available information indicates that 2.1 litersof expired alveolar air contain approximately the same quantity of alcohol as 1 milliliter of blood". [n.63] Again the "1:2100" model appears to have gathered simply a consensus for its possibility. In fact, this uncertainty in the blood alcohol:breath alcohol ratio led some noted authorities to suggest, as *195 early as 1974, that attempts to characterize an exact ratio for all individuals be abandoned. [n.64]

Out of the starting gate the breathalyzer was empirically validated for reliably predicting the ingestion of alcohol. But it was not, and still remains, unvalidated in the ability to predict an accurate blood alcohol content for all human beings owing the unvalidated "1:2100" model. At best, a consensus exists only for the possibility of such a model. This should have never been construed as a consensus worthy of elevating the breathalyzer results through some legal test of admissibility at a civil or criminal trial on the disputed issue of blood alcohol content. Although the Nebraska Supreme Court ultimately recognized the scope of the empirical verification behind the "1:2100" model, courts must begin to scope verification of a newly postulated model as an integral part of any admissability criterion.

B. The Battered Woman's Syndrome

It would be useful to consider the earlier suggestion that history is truly one of the borderlands marching between scientific and non-scientific pursuits. To start, no substantial general principle of historical explanations have won universal acceptance. Whatever their reasons, historians do not agree on the general theoretical foundations of the methodology of their studies. Instead of establishing, by mutual criticism and tacit cooperation, a limited common basis of acceptable models from which to build upwards and outwards, they often feel bound to set up antagonistic "schools" of interpretation. [n.65]

Of course, history is not usually the subject of an admissability question. Under our understanding of legal consensibility (Frye) it would be safe to conclude that expert testimony explaining some particular event would not be generally accepted by all historians. Alternatively, under our scientific understanding of consensibility, we would note a lack of empirical verification, no cooperation towards a limited common basis of acceptable historical models and the presence of many unsolved problems in explaining why a certain particular event occurs.

*196 But disciplines such as sociology and social psychology have been recently offered, in some cases admitted into evidence, but to a large extent they raise the same problems which have been described above concerning history. The point here is that as a consequence of applying a purely legal standard of admissability, which ignores any of the demarcations that have now been exposed concerning the scientific understanding of consensibility, non- science has made its way into the jury room.

For example, a large amount of literature on both sides, has emerged, concerning the reliability of the battered woman syndrome defense. [n.66] Nevertheless, as of the Fall of 1982 there were 50 cases around the country in which lawyers had attempted to get expert testimony on the syndrome admitted. The testimony was excluded in only four cases, and two of those four cases were overturned on appeal. [n.67] In short, such expert testimony is offered on the issue of the existence and reasonableness of the accused's belief that force was necessary to protect herself from imminent harm. [n.68]

The battered woman's syndrome had its origin through the work of Dr. Leonore Walker, a self-described feminist psychologist at the Colorado Women's College in Denver.

[n.69] At the core of the battered woman's syndrome lie the theories of "learned helplessness" and the "cycle of violence". [n.70] Learned helplessness is a psychosocial theory for lack of response, or passive behavior in the face of the ability to act. [n.71] The learned helplessness theory thus explains why the battered woman who kills has not left the situation prior to killing. The cycle theory of violence identifies three components of a battering relationship varying in intensity and time: (1) the tension building phase; (2) the explosion or acute battering incident; and (3) the calm, loving respite. [n.72]

*197 In State v. Kelly, the New Jersey Supreme Court held that expert testimony concerning the battered woman syndrome was admissable [n.73] The court ruled that the testimony was relevant under New Jersey's standard of self-defense, and that the testimony met the standards of New Jersey's rules for admissability of expert testimony. The New Jersey standard contemplated a consensus idea in that the field to be testified to must be at a state of the art such that an expert's testimony could be sufficiently reliable. [n.74]

An amicus curiae brief on behalf on the American Psychological Association was submitted to the New Jersey Supreme Court in support of admissability which focused on the issue of scientific reliability. In the argument for admissability of expert testimony concerning battered woman, the brief argued that the relevant case law, the scientific literature, the law review commentary, and the number of researchers and scientific materials in the field-supported expert testimony on the study of battered woman. [n.75]

In an amicus curiae brief filed by The American Civil Liberties Union of New Jersey, Dr. Leonore Walker, in an accompanying affidavit, cited her own two theories which described the psychology of battered woman ("learned helplessness" and "cycle of violence"). [n.76] Dr. Walker also suggested that the methodology utilized to establish such theories was approved by the relevant scientific community. [n.77] For example, studies in over 400 battered woman confirmed the theory of "cycle of violence". [n.78] Through methods of in-depth interviewing it was also established that most battered woman develop a unique coping style ("learned helplessness"). [n.79]

*198 Dr. Walker then concluded that such information on battered woman would be useful to understand how the effects of previous abuse impact upon a woman's state of mind during some incident at question so that it was reasonable for her to believe that the violence would escalate to life- threatening proportions. [n.80] Consequently, such expert testimony would be critical to support a self-defense plea.

Under a legal standard of "general acceptance" one cannot criticize the Kelly court for ruling as they did. The literature contained at least 5 books and 70 scientific papers about the battered woman's syndrome. [n.81] Other jurisdictions were on record as having accepted such testimony. Dr. Walker's methods appeared to be accepted by the relevant scientific community and her theories were supported by the American Psychological Association.

But under our scientific understanding of modelling, verifiability and consensibility, the Kelly opinion would come out differently. There appears to be two operative models. One predicts why a battered woman does not leave subsequent to battering. Consequently, the model of "learned helplessness" goes to prediction of lack of response by the battered woman. The model of "cycle theory" goes to a prediction of future beatings, repetitive beatings, on the part of the male partner. [n.82]

Dr. Walker has noted that in ninety-six cases of battered woman who have killed or seriously hurt their abusers, in self-defense, she was successful in introducing evidence of battered women syndrome at the trial of thirty-three cases and in the sentencing phase of thirty-two cases. [n.83] Since self- defense focuses on the act of defending oneself, it rests on a determination that the act was right because of its circumstances. [n.84] At common law, a person who kills another pursuant to an honest and reasonable, albeit mistaken, belief that such act was necessary to preserve their life is not guilty of any homicidal offense. [n.85] But neither of Dr. Walker's two models predict that when a woman is confronted *199 with a repeated battering circumstance she would likely perceive a need to preserve life (therefore a reasonable belief) in a subsequent non-confrontational episode, and kill. Dr. Walker's models only predict why the woman do not leave the battering spouse, as well as the cyclical nature of the beatings. From this latter perspective, the male's behavior has been much of the focus of predictability and verification.

Until Dr. Walker can predict that under conditions of battering the actor perceives a need to preserve life in a subsequent non-confrontational setting, and kills or will kill in self-defense, her testimony on an issue of self- defense should not be admitted. In other words, there needs to be a model which focuses on the predictability of the woman to resort to deadly force in an act of self-defense, when battered and then placed outside of a violent episode. [n.86] The scientific method, postulating a model, and providing empirical verification, would demand nothing less. The Ohio Supreme Court, in rejecting the battered woman syndrome testimony outright stated, "Expert testimony on the "battered wife syndrome" ... to support defendant's claim of self-defense is inadmissable herein because (1) it is irrelevant and immaterial to the issue of whether the defendant acted in self-defense at the time of the shooting ...". [n.87] Although the Ohio Supreme Court did not recognize the issue as recited herein, their perception contemplates the failed nexus between existing models and a claim for justification.

Under a scientific criterion of consensibility, expert testimony on the battered woman's syndrome would also be excluded. Recall the three-prong test proposed by Ziman. [n.88] First up is whether or not consensus through rejection or acceptance of empirical testing has been the main objective. There has been no empirical testing of the "preserve life model" and one can conclude that consensus through empirical testing has not been a main objective.

*200 Secondly, one must consider whether or not there has been mutual criticism and tacit cooperation towards a limited common basis of acceptable models. Note Dr. Walker's own comments from her original work, "As a trained researcher, I felt uneasy stating some of my conclusions in this book. They seemed too tentative to write down in the positive manner which I have used." [n.89] It is difficult to characterize this perception as aneffort to develop a common basis of acceptable models.

As the final consideration by Ziman, non-science is established when there has been legitimate reason for widespread intolerance of the views of one scholar by another. Apart from the split of commentary in the literature, [n.90] consider the position of Walker's supporters, "Although Dr. Walker did not use a control group for purposes of comparison this was not necessary to insure a scientifically valid study ... it was perfectly legitimate ... not to use ... a control group for purposes for comparison". [n.91] An in-depth interview technique was argued as the alternative. [n.92] The psychological community, while accepting such interviewing techniques of Dr. Walker, discards conclusions (i.e., models) that are not based on random sampling. [n.93] It seems quite legitimate to reject Dr. Walker's views since her methodology did not include random sampling. Dr. Walker by her own account recognized that since women were not randomly selected, they could not be considered a legitimate data base from which to make specific generalizations. [n.94]

In sum, the Ziman test begins with the notion that if we are to rely on consensus, we must consider why that has been the case. And if we discover that consensus through empirical testing has not been the main objective, that development of a few empirically verified models has been avoided, it is no doubt a warning sign. When we add to this the fact that legitimate reasons exist for a widespread difference of opinion on the validity of a model, a non- scientific body of information is present making predictions that may be no better than pure chance. While the court in Kelly may have been impressed that there was a great deal of *201 literature on the subject of battered woman's syndrome, we should consider whether any portion of such literature identifies the proposed model, is characterized by empirical validity, and is making a prediction relevant to some self-defense plea. Absent empirical validity, the consensual support should be studied carefully, along the lines suggested by Ziman. Upon this analysis it appears that Walker's testimony on the battered woman's syndrome would not be admissable as scientific in the context of some reliability to predict behavior in a nonconfrontational setting.

Conclusions--A Scientific Standard of Admissability

The drafters of the Federal Rules of Evidence could not reach agreement on the proper role of the Frye standard for the admissability of scientific evidence. Over 65 years of common law development has centered on some sort of consensual basis to decide whether or not a given piece of scientific evidence should be admitted. Recent judicial attempts to resolve the debate over Frye have diluted, rather than eliminate, the role of consensus in determining reliability of scientific evidence (Downing).

By considering what the scientific community judges as being good or bad science, we have learned that science is a method of prediction through the use of empirically verified models. Typically models serve to connect present events with future events and utilize the knowledge of prediction to shape future physical events as they occur. Deferring to this understanding of science suggests that evidence should be labelled scientific when it is offered to make a prediction of fact.

Models do not become verified as a consequence of social powers (consensus). Rather, models become verified when their predictions are subjected to the intensity of repeated and rigidly controlled experiments. Courts should fully subordinate the notion that consensus has an important role in establishing scientific validity.

A scientific standard of admissability would require identification of the postulated model and the type of prediction offered. Subsequently, the scope of verification would be considered to ascertain the level of empirical verification. In other words, the reliability of the prediction and its relevance should become central to the admissability issue. Instinctively we would reject predictions that are no better than pure chance. Alternatively, predictions that are fully verified would have enormous value in the fact-finding process.

Short of empirical verification, consensual support must be studied carefully. Scientists have long recognized that the idea of consensus goes further than a simple accounting of the number of supporters for a given *202 model. On an initial level, consensus that represents some indicia of response to empirical testing may be useful in determining validity. Consensus that stands apart from empirical testing is an early sign of a non-scientific body of information. Courts must reject the introduction of predicted facts simply because some number of people believe that such predictions are accurate.

Following this proposal it would be naive to overlook the argument that courts may be confronted with extraordinary amounts of information every time some expert expressed an intent to offer scientific evidence. Nevertheless, and in the case of the availability of more and more scientific evidence, we cannot discount the obligation, under the laws of evidence, to filter non- science from the fact-finder. That is, in an age dominated by the works of science and technology, law and science must and will interact and work together. The issue is not "whether" but "how". [n.95]

[n.a]. Associate Professor of Engineering, University of Lowell, Lowell Massachusetts. B.S., University of Connecticut, Storrs Connecticut. Ph.D., University of Massachusetts, Amherst. J.D., Franklin Pierce Law Center, Concord, New Hampshire. Associate, Hayes Soloway, Hennessey & Hage, Manchester, New Hampshire.

[n.1]. Markey, Science and Law: The Friendly Enemies, 30 Idea 13, 18 (1989).

[n.2]. Doyle, Applying Lawyers' Expertise To Scientific Experts: Some Thoughts About Trial Court Analysis Of The Prejudicial Effects of Admitting And Excluding Expert Scientific Testimony, 25 Wm. & Mary L.Rev. 619, 620 (1984).

[n.3]. See, e.g., S. Jasanoff, Science of the Witness Stand, Issues in Sci and Tech., Fall 1989 at 80.

"Even when court seem to defer to the authority of science, the construction of knowledge in the courtroom usually proceeds in accordance with legal, rather than scientific rules. In the process, judges may draw conclusions that are unacceptable or incomprehensible to scientists but that reaffirm the separate, even superior authority of the legal system."

Id. at 87.

[n.4]. A related and unsettled issue concerning scientific evidence has focused on the question of judicial capacity to understand scientific evidence. See Wesley, The Question of Judicial Capacity, 25 Wm. & Mary L.Rev. 675 (1984); Blair, The Court of Appeals for the Federal Circuit--Should Its Judges Be Technologically Literate or Illiterate?, 27 Idea 121 (1986); Martin, The Proposed "Science Court", 75 Mich.L.Rev. 1058 (1977). While this present study proposes a new standard of admissability for scientific data, it remains whether the greater use of scientific evidence in the courtroom threatens the ability of judges and jurors to comprehend such testimony.

[n.5]. Frye v. United States, 293 F. 1013 (D.C.Cir.1923).

[n.6]. Tenney, The Horizontal Gaze Nystagmus Test and The Admissability of Scientific Evidence, 27 N.H.Bar Journal 179 (1986).

[n.7]. Note 5 supra at 1014.

[n.8]. United States v. Williams, 583 F.2d 1194, 1198 (1978).

[n.9]. Wesley, Scientific Evidence and the Question of Judicial Capacity, 25 Wm. & Mary L.Rev. 675, 684 (1984).

[n.10]. G. LILLY, AN INTRODUCTION TO THE LAW OF EVIDENCE, 494 (1987).

[n.11]. See e.g., Scientific Evidence Symposium, 25 Wm. & Mary L.Rev. 517-705 (1984).

[n.12]. United States v. Downing, 753 F.2d (3d Cir.1985).

[n.13]. State v. Coolidge, 109 N.H. 403, 260 A.2d 547 (1969).

[n.14]. Gianelli, Background Paper Prepared for the National Conference of Lawyers and Scientists, 99 F.R.D. 189, 198 (1983).

[n.15]. Surprisingly, these terms have been used interchangeably by many courts, ignoring a distinction well known in the scientific community. See infra, note 27 and accompanying text.

[n.16]. Note 8 supra, at 1198.

[n.17]. See Coppolino v. State, 223 So.2d 68 (Fla.App.1968) (the court upheld the admissability of the results of a relatively unknown test on the grounds that novel tests are not necessarily inadmissable simply because the profession at large is not yet familiar with them).

[n.18]. Lacey, Scientific Evidence, 24 Jurimetrics J. 254, 265 (1984).

[n.19]. State v. Catanese, 368 So.2d 975 (La.1979); State v. Walstad, 119 Wis.2d 483; 351 N.W.2d 469 (1984).

[n.20]. The development and enactment of the Federal Rules of Evidence in 1975 did not explicitly mention the Frye standard. Some commentators have argued that the silence of the Federal Rules and its drafters should be regarded as tantamount to an abandonment of the general acceptance standard (Frye). 3 J. WEINSTEIN & M. BERGER, WEINSTEIN'S EVIDENCE 702[02], at 702- 716 (1982). Others have suggested that since Frye was the established general rule and no statement repudiating Frye appears in the legislative history, the general acceptance standard remains intact. Giannelli, The Admissability of Novel Scientific Evidence: Frye v. United States, A Half- Century Later, 80 Colum.L.Rev. 1197, 1229 (1980). Perhaps the only safe comment to the silence of the Federal Rules on the issue of Frye is that the drafters could not reach a conclusion as to the general acceptance standard and opted for additional common law development.

[n.21]. Rule 702 of the Federal Rules of Evidence provides that expert testimony is admissable "[i]f scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue." Fed.R.Evid. 702.

[n.22]. 753 F.2d 1224, 1237 (3rd Cir.1985).

[n.23]. Id. at 1237.

[n.24]. Id.

[n.25]. Id. at 1238, 1239.

[n.26]. Id. at 1239.

[n.27]. It is worth noting at this point that the Downing court has touched on an important distinction that might have been worth developing in more detail. Namely, the court's definition of reliability (recall, the frequency with which a technique leads to erroneous results) approaches what is known in the scientific community as precision. That is, only if an observation is repeated a number of times and only if the same result is obtained each time is it possible to be reasonably sure that the observation is precise. Alternatively, accuracy is reserved for the implications that are associated with some precise observation. For example, while one may repeat a similar observation a number of times with the same result, one may not be sure of what one is measuring. That may sound silly, but it is a very real problem. It may be illustrated by one commonly used method of chemical analysis. Many measurements of chemical concentration depend on the fact that the substance under study reacts with another test chemical to give a particular color. Suppose that one desires to measure the concentration in blood of substance X. It is known that when X reacts with a test chemical Y, a blue color develops whose precise intensity can be measured by means of an instrument known as a colorimeter. The higher the concentration of X, the more intense is the color which develops. By measuring the intensity of the color it is therefore possible to estimate the concentration of X in the blood. It is apparent that if this method of measurement is to be accurate, one important condition must be fulfilled. There must be no other substance in the blood apart from X which reacts with Y to give a color. If some other substance is present, the result, although precise and reproducible, is meaningless. The literature of biological science is littered with false observations of scientists who have measured not only what they thought they were measuring, but a number of other extraneous substances as well. See D. HORROBIN, SCIENCE IS GOD, at 68 (1969). This discussion is developed more fully at note 43 and accompanying text.

[n.28]. 753 F.2d at 1242.

[n.29]. Id. Note that this requirement insures that the trial court's decision on admissability can be effectively reviewed on appeal.

[n.30]. See, ZIMAN, What Is Science, in INTRODUCTORY READINGS IN THE PHILOSOPHY OF SCIENCE 35 (E. Klemke, R. Hollinger, A. Kline ed. 1980). See also, J. KEMENY, A PHILOSOPHER LOOKS AT SCIENCE 174 (1959).

[n.31]. M. WALKER, THE NATURE OF SCIENTIFIC THOUGHT (1963); C. LASTRUCCI, THE SCIENTIFIC APPROACH, BASIC PRINCIPLES OF THE SCIENTIFIC METHOD (1967).

[n.32]. See LASTRUCCI, Id. at 235.

[n.33]. For example, the statement, "it is raining outside," may be labeled as true or false by looking outside. It is capable of 100% verification.

[n.34]. Moeller v. St. Paul City R. Co. 218 Minn 353, 16 NW2d 289 (1944) holding that expert evidence as to the rate of speed of a vehicle was admissible although based on facts other than personal observation.

[n.35]. See generally, ERWIN, DEFENDING DRUNK DRIVING CASES, § § 14.01-28.06 (3d. ed. Rev.1982).

[n.36]. See LILLY, note 10 supra at 507.

[n.37]. Id. at 504.

[n.38]. 31 AM.JUR.2D Expert and Opinion Testimony § 128 (1967).

[n.39]. In the United States, 0-45% of the population, A-41% of the population, B-10% of the population and AB-4% of the population.

[n.40]. See LILLY, note 10 supra at 500.

[n.41]. See LASTRUCCI, note 31 supra at 235, from which this discussion closely follows.

[n.42]. The model of radio waves changing frequency upon striking a moving object is open to this type of challenge. Consider a police officer who inadvertently moves his radar gun upon the approach of traffic. This results in an alteration in the state of the object being observed in so far as the approaching vehicle has been accelerated by the act of the officer's unsteady hand. See the discussion above that follows this note.

[n.43]. Our model of blood typing is open to this sort of challenge. Consider that the experimental technique for classifying blood types turns on the detection of certain biochemical compounds in our bloodstream. If the measurements used to detect these certain biochemical compounds respond to the presence of other components in the blood, an erroneous classification could be made. This is not to suggest that our current methods for measuring blood typing is in doubt. It is made simply so that one can appreciate the importance of the notion that science is always concerned about the possibility that one is not measuring what was intended to be measured.

[n.44]. See J. KEMENY, note 30, supra at 79.

[n.45]. Prof. Horrobin articulates a distinction between bias and dishonesty. Whereas dishonesty is a deliberate attempt to falsify some results, bias is considered an unconscious falsification which creeps in because the observer thinks he knows what the results should be. See HORROBIN, note 27 supra, at 69.

[n.46]. See POPPER, Science: Conjectures and Refutations, in INTRODUCTORY READINGS IN THE PHILOSOPHY OF SCIENCE 19 (E. Klemke, R. Hollinger, A. Kline ed. 1980).

[n.47]. Jacobellis v. Ohio, 378 U.S. 184, 197 (1964).

[n.48]. MILTON R. WESSEL, SCIENCE AND CONSCIENCE 143 (1980). Prof. Wessel states the complicated nature of scientific consensus in the following manner:

"A scientific consensus can be independent of the underlying opinion to which the consensus relates. Thus, scientists may agree that there is a consensus regarding a scientific opinion, even though some take exception to that opinion (after all, it was once the consensus that the earth was flat and that man sprang from Adam). Similarly, they may be in complete agreement about the underlying opinion and yet conclude that there is not yet a scientific consensus about it (for example, because of inadequate scientific participation, or insufficient time for review). Scientific consensus is therefore very different from scientific opinion. Moreover, and perhaps surprisingly, there will far more often be complete unanimity regarding the existence or lack of consensus, then there is with regard to the underlying opinion. Where consensus exists, it is enormously credible and persuasive. Whatever consensus is, scientists alone know it when they see it. Accordingly, any scientific consensus-finding mechanism should itself determine the extent to which it has found consensus and then communicate that fact to the public."

[n.49]. See, WALKER, note 31 at 13.

[n.50]. See, ZIMAN, note 30 supra at 53. This criterion of scientific consensibility is later characterized as a three-prong test. See infra, note 88 and accompanying text.

[n.51]. See, P.R. THAGARD, Why Astrology Is A Pseudoscience, in INTRODUCTORY READINGS IN THE PHILOSOPHY OF SCIENCE 70 (E. Klemke, R. Hollinger, A. Kline ed. 1980).

[n.52]. This demarcation of disciplines as non-science is not my own. See H. JONES, Legal Inquiry and the Methods of Science, in LAW AND THE SOCIAL ROLE OF SCIENCE 120, 127 (H. Jones, ed. 1967); ZIMAN, note 30 supra, at 52.

[n.53]. H. JONES, Id. at 128.

[n.54]. See B. BUENO DE MESQUITA, D. NEWMAN, A. RABUSHKA, FORECASTING POLITICAL EVENTS: THE FUTURE OF HONG KONG (1985).

[n.55]. A. MOENSSENS, F. INBAU, J. STARRS, SCIENTIFIC EVIDENCE IN CRIMINAL CASES 79 (1986), from which this discussion closely follows.

[n.56]. Id. at 88.

[n.57]. ERWIN, note 35 supra at § 18.01.

[n.58]. Most states have statutes providing for chemical tests for intoxication in traffic related incidents. In most chemical test statutes, the following presumptions are prescribed:

1. A subject whose blood alcohol content was less than 0.05% was presumed not to be under the influence.
2. Where the alcohol content was in excess of 0.05% but less than 0.10% there was no presumption either way.
3. A level of 0.10% or higher gave rise to a presumption of being under the influence.

Such statutory presumptions embrace both of the postulated models. See A. MOENSSENS, F. INBAU, J. STARRS, note 55 supra at 94.

[n.59]. Id. at § 18.02. Erwin notes that recent studies establish the range of blood alcohol: breath alcohol ratios from 1706 to 3063. Consider that men will generally have more alcohol dissolved in the vapor above their blood as compared to woman.

[n.60]. 224 Neb. 725, 400 N.W.2d 872 (1987). While this may appear to be a compromise solution, note that this reduction only considers the possible range of blood alcohol:breath alcohol ratios. Other factors could disrupt the correlation further, e.g., contamination of breath by mouth alcohol, alcohol brought up by the stomach, foreign objects in the mouth, and alcohol retention by dentures and denture adhesives. See ERWIN, note 35 supra at § 18.02.

[n.61]. ERWIN, note 35 supra at § 18.01.

[n.62]. Id.

[n.63]. Id.

[n.64]. M.F. Mason and K.M. Dubowski, Alcohol, Traffic, and Chemical Testing in the United States: A Resume and Some Remaining Problems, 20 Clinical Chemistry 2 at 126-140 (1974).

[n.65]. ZIMM, note 30 supra at 47. Prof. Walker adds that history, by definition, carries some implication of predictive use, but most historians have not attempted to predict explicitly the future of nations and cultures. Although models of history designed for predictive use have been constructed by Hegel, Marx, Spengler, and Toynbee, they permit only long-range, general predictions, and not enough time has yet elapsed to permit validation or invalidation of any of these models. WALKER, note 31 supra at 150.

[n.66]. Favorable: v. Mather, The Skeleton in the Closet: The Battered Woman Syndrome, Self-Defense and Expert Testimony, 39 Mercer L.Rev. 545 (1988); Skeptical: M. Mihajlovich, Does Plight Make Right: The Battered Woman Syndrome, Expert Testimony and the Law of Self-Defense, 62 Ind.L.J. 1263 (1987).

[n.67]. E. Schneider, Describing and Changing: Woman's Self Defense Work and the Problem of Expert Testimony on Battering, 9 Women's Rights L.Rptr. 195, 200 (1986).

[n.68]. State v. Leidholm, 334 N.W.2d 811, 820 (N.D.1983); State v. Allery, 597 P.2d 312, 316 (1984) (Expert testimony would allow the defense "[t]o effectively present the situation as perceived by the defendant, and the reasonableness of her fear".).

[n.69]. M. Mihajlovich, note 66 supra at 1257.

[n.70]. L. WALKER, THE BATTERED WOMAN (1979).

[n.71]. Id. at 47.

[n.72]. Id. at 55.

[n.73]. State v. Kelly, 478 A.2d 364, 373, (1984). The New Jersey Supreme Court noted that while the record before them revealed that the battered woman's syndrome had a sufficiently reliable basis to produce uniform and reasonably reliable results, the case would be remanded for a new trial so that the State could cross-examine the expert on methodology and assertion that the battered woman's syndrome had been accepted by the relevant scientific community. Id. at 381. One judge dissented from the remand on the issue of scientific acceptability, finding that the defendant had sufficiently demonstrated both scientific acceptance and reliability. Id. at 385 (Handler, J., concurring in part and dissenting in part).

[n.74]. Brief of Amicus Curiae, American Psychological Association in Support of Appellant, State v. Kelly, 478 A.2d 364 (1984). This brief is reproduced at 9 Women's Rights L.Rptr. 253 (1986).

[n.75]. See, 9 Women's Right L.Rptr. at 254 (1986).

[n.76]. Id. at 245, 250, 251.

[n.77]. Id. at 251.

[n.78]. Id.

[n.79]. Id.

[n.80]. Id. But whether or not most battered woman reasonably believed that violence would escalate to life-threatening proportions and killing their spouse was the only solution, in a non-confrontational setting, remains the key question for further research. See infra, note 86 and accompanying text.

[n.81]. State v. Kelly, 478 A.2d at 380.

[n.82]. See Walker, note 70 supra.

[n.83]. L. Walker, A Response to Elizabeth M. Schneider's Describing and Changing: Women's Self-Defense Work and the Problem of Expert Testimony on Battering, 9 Women's Rights L.Rptr. 223, 224 (1986).

[n.84]. Schneider, note 67 supra at 215.

[n.85]. W. LAFAVE & A. SCOTT, CRIMINAL LAW 454 (2d. ed. 1986).

[n.86]. See M. Mihajlovich, note 66 supra at 1270-1271.

"Self-defense in the common plea of battered woman who kill their abusers during a violent episode, in anticipation of a violent episode, after a violent episode, and while the

abuser is asleep or incapacitated. Proponents attempt to introduce expert testimony in all of these cases (citation omitted). Where the killing occurs during a violent episode, self-defense is an appropriate plea which needs little adjustment by expert testimony at trial. But expert testimony is vital when the defendant kills her batterer outside a violent episode. The two areas of self-defense requiring expert opinion are the reasonableness of the degree of force used and the reasonableness of the battered woman's perception of danger. In the non-confrontational killings, by battered woman, expert testimony attempts to create these elements of self-defense where they do not exist (emphasis added)."
Id. at 1270-1271.

[n.87]. State v. Thomas 423 N.E.2d 137 (1981).

[n.88]. See, ZIMAN, note 30 supra at 53.

[n.89]. L. Walker, supra note 70 at xv-xvi. At least two courts have looked to this language in their decisions to limit expert testimony on the battered woman's syndrome. See Ibn-Tamas, 455 A.2d 893, 894 (D.C.1983); Buhrle v. State, 627 P.2d 1374, 1377 (Wyo.1981).

[n.90]. See note 66 supra.

[n.91]. Brief of Amicus Curiae, The American Civil Liberties Union of New Jersey, Affidavit of S.K. Steinmetz, State v. Kelly, 478 A.2d 364 (1984). The brief is reproduced at 9 Women's Rights L.Rptr. at 254 (1986).

[n.92]. Id.

[n.93]. See Mihajlovich, note 66 supra at 1267.

[n.94]. See Walker, note 70 supra at xiii.

[n.95]. Loevinger, Jurimetrics: Science in Law, in SCIENTISTS IN THE LEGAL SYSTEM 22 (W. Thomas ed. 1974).