Forensic experts' perceptions of expert bias

Michael Lamport Commons [a,*], Patrice Marie Miller [a], Eva Yujia Li [b], Thomas Gordon Gutheil [a]

[a] Department of Psychiatry, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA
[b] Dare Institute, Cambridge, MA, USA

ARTICLE INFO

ABSTRACT

How do expert witnesses perceive the possible biases of their fellow expert witnesses? Participants, who were attendees at a workshop at the American Academy of Psychiatry and Law were asked to rate for their biasing potential a number of situations that might affect the behavior of an opposing expert. A Rasch analysis produced a linear scale as to the perceived biasing potential of these different kinds of situations from the most biasing to the least biasing. Working for only one side in both civil and criminal cases had large scaled values and also were the first factor. In interesting contrast, a) an opposing expert also serving as the litigant's treater and b) an opposing expert being viewed as a "hired gun" (supplying an opinion only for money) were two situations viewed as not very biasing. Order of Hierarchical Complexity also accounted for items from the 1st, 2nd and 3rd factors. The result suggests that the difficulty in understanding the conceptual basis of bias underlies the perception of how biased a behavior or a situation is. The more difficult to understand the questionnaire item, the less biasing its behavior or situation is perceived by participants.

© 2012 Published by Elsevier Ltd.

1. Introduction

The objectivity that an expert witness brings to the legal system is the most valued quality of an expert, whether that expertise is in psychiatry or elsewhere. One of the most challenging but necessary ideals for expert witnesses to uphold, therefore, is dealing with "expert bias." Expert bias is seen as a deviation from the "ideal" neutral balanced assessments, judgments and the like. Perceived bias here is operationally defined as how strongly biasing the study participants found certain situations to be. Participants in the study were asked to rate on a scale of 1 to 6 how potentially biasing are situations as described in the items, or how biased they believe the expert witness as described in the item is. Here we concentrate on the perceptions of expert witnesses themselves because of the large degree of experience they have in observing possibly biasing situations. We differentiate between experts' perceptions and juror and judges' perceptions, making no assumptions about their similarity or differences.

In our previous work on expert bias (Commons, Miller, & Gutheil, 2004), we showed that expert witnesses in our survey perceived the existence of a good deal of such bias. In that previous study, it appeared as if some situations were perceived as more biasing than others. In other words, some of the potentially biasing situations had higher significance values and larger effect sizes. This did allow a conclusion that some situations were perceived by experts to be actually more biasing than others, but there was no way to ascertain specifically how much more or how much less biasing each situation was perceived to be. The purpose of the current study is to find out how potentially biasing each of the situations is perceived to be, on a ruler-like scale with the numbers on the ruler being equally spaced. Jurors understand measures that are ruler-like because they understand measurement with rulers. Hence, jurors may readily understand the results of using a technique called Rasch analysis (Rasch, 1960) as described below. We hoped that forensic experts might benefit from being informed as to the perceived degree of seriousness of various biasing situations. With such information, forensic experts can consider altering their own behavior and/or informing the jury of the seriousness of biases which the other side may hold in a cross examination.

The second focus of the current study is to understand perceived bias from a psychophysical perspective. In psychophysics, one finds the properties of stimuli to explain the responses. In this study, we can predict the perceived bias of a questionnaire item by noting the subject's difficulty in understanding that item. Past research has explained why and how bias exists: cognitive shortcuts such as heuristics, availability bias, confirmation bias, and self-serving bias whose goal is to protect self-esteem and support optimism in people, and so forth (Kahneman, Slovic, & Tversky, 1982). A question that remains is why some biases are easier to identify and overcome than others. For example, it is obvious that having monetary interest in the outcome of a case is a biasing factor. In order to eliminate the effect of such bias on decision making, an interest-free, third party is usually called upon to make judgment of a conflict between two opposing sides. It is less obvious, for example, to realize that absolute

* Corresponding author at: Department of Psychiatry, Beth Israel Deaconess Medical Center, Harvard Medical School, 234 Huron Avenue, Cambridge, MA 02138, USA. Tel.: +1 617 497 5270.

E-mail address: commons@tiac.net (M.L. Commons).

bias-freeness is impossible, and holding such a belief is itself an indication of bias. This paper explores the possibility that the difficulty of understanding the items may influence how expert witnesses rate the biasing potential of a situation or the degree of bias of the subject.

One underlying factor that may predict perceived bias is the difficulty of the task. It is assumed that tasks have an inherent difficulty. Here that difficulty is called *Order of Hierarchical Complexity* (Commons, 2008; Commons, Trudeau, Stein, Richards, & Krause, 1998). The tasks in our study are reasoning about how biasing a given situation is. We predict that the more hierarchically complex the task, the less likely people are going to perceive bias.

1.1. Rasch analysis

In this study, Rasch analysis is used to show the degree of perceived bias in an objective, empirical manner. In order to understand our results, some basic knowledge of Rasch scales is necessary. A Rasch model produces an additive scale. It can be used to analyze a large variety of human sciences data. This model, for example, through the use of probabilistic equations, converts raw ratings of items into scales of Rasch scores that have equal intervals. Such a scale can then be used as a type of ruler against which to measure the data on survey items as well as on respondents (Bond & Fox, 2007). Statistically speaking, this scale will be linear (Bradley & Terry, 1952; Luce, 1959). As a result, a change of severity of the Rasch scores of 1, is the same going from -2 to -1 as going from 0 to $+1$. Doubling on the Rasch scale means the same change in severity anywhere along its linear axis. For example, a perceived bias with a value of 2.3 is half as severe as a perceived bias of 4.6. After analyzing data with a Rasch model, a number of questions can be answered. First, where on the scale does each independent variable fall (e.g. in this case, how severe are each of the perceived biases for an item)? Second, what is the range of scaled values between all variables for all participants? Third, for each participant, what is the scaled value of overall severity of these biases when measured on the same scale?

The answer to the first question will give expert witnesses a scale of how biasing each situation appears. This should allow experts to point to some empirical data when confronted with some of these situations. The answer to the second question should help answer: How much of a difference does a change in score make? Consider the difference of 1 unit between two scores, for example an item with a score of 1.5 and an item of 2.5. For a small range of scaled perceived bias scores, this would be a big difference, whereas for a large range this would be a small difference.

1.2. Factor analysis

Another way to consider the issue of how bias is perceived is to examine characteristics of the items themselves. Although the data might fit a one dimensional Rasch model, there can still be characteristics of the items that form sub-dimensions. In order to explore sub-dimensions of items, a factor analysis will be performed. Factor analysis is a method to explain correlations between observed variables, in this case, items in the questionnaire. It uncovers unobserved, latent properties of the items called factors (Gorsuch, 1983).

1.3. Difficulty of items as measured by Model of Hierarchical Complexity

Next, we explore the a priori difficulty of understanding the biasing potential of an item. What is its relationship to the factors found from the factor analysis? To explain the factors, we propose to consider how these items are viewed: what is the required stage of performance needed successfully to understand an item (Commons, 2008; Commons & Miller, 1998)? More difficult items, requiring higher stages of conceptual development, may be found to be less biasing.

The Model of Hierarchical Complexity (MHC) measures the a-priori difficulty of tasks. Adults vary greatly in how detailed a task they understand and do. Because less complex tasks must be completed and practiced before more complex tasks can be acquired, the Model argues that this accounts for the developmental changes seen in individual persons' performance on tasks. For example, persons cannot perform arithmetic until they can truly and correctly count. In order for difficulty to be precisely measured, the Model proposes a metric. That is, that Task A is considered to be hierarchically more difficult or complex than Task B if Task A is made up of two or more simpler actions (such as Task B and a third task, C), and these simpler task actions are coordinated in a non-arbitrary way. If Task A consisted of a combination of Task B and Task C, then it would be what is called one Order of Complexity higher than Tasks B and C. The Model specifies that there are 16 orders of complexity, starting with tasks that are completed by the simplest animals and infants, and progressing to highly complex tasks that only some adults complete. These orders are shown in Table 1.

An individual's stage of development or performance is based on the Order of Hierarchical Complexity of the task that he or she correctly completes, and because of that is given the same name and number as the Order of Hierarchical Complexity of the task. So, if an individual completes a task that is at order 10 (Formal), performance on that task is also considered to be at the Formal Stage.

The Model of Hierarchical Complexity (MHC) has been shown to account for performances in a variety of different domains (Commons, 1999). For example, when the action described by an item is at a higher order than a task that the participant can understand, the grasp of the value of the task is too high for the participant, and the participant cannot perceive its value.

The order of complexity of the task is determined through analyzing the demands of each task by breaking it down into its constituent parts. The following is a list of tasks people do at each Order of Hierarchical Complexity from 8 to 12. At each order, key features are described and then examples of such tasks are given. They should be understood as only examples but not an exhaustive list. Tasks of all domains can theoretically be mapped to this scale.

At the *concrete order 8*, two or more primary stage 7 operations may be coordinated. Coordinating two perspectives becomes possible and deals can be made. People respond to threats by making a deal. For example, the insurance company lawyer says to an expert, "if you do not say what I want you to say, you will never work again in this town." Giving in to such a threat creates bias. However, negotiations are specific to the person that one is dealing with and based on concrete experiences.

At the *abstract order 9*, two or more concrete order 8 operations may be coordinated. It becomes possible to coordinate concrete instances and form the notion of a variable and understand the value of the variable. For example, concrete interactions with people may lead to the understanding of social norm. People may figure out what their responsibilities are on a job, based on what the socially accepted role of the position is. For example, an expert witness may know that the social norm of this position is to be bias-free. People performing at this stage have an idea of a variable, such as acting as

Table 1
Orders of Hierarchical Complexity.

Order	Name	Order	Name
0	Calculatory	8	Concrete
1	Sensory or motor	9	Abstract
2	Circular sensory-motor	10	Formal
3	Sensory-motor	11	Systematic
4	Nominal	12	Metasystematic
5	Sentential	13	Paradigmatic
6	Preoperational	14	Cross-paradigmatic
7	Primary	15	Metacrossparadigmatic

a consultant or acting as an expert. However, they have no logical or empirical way of deciding if some activity met the criteria for each role.

At the *formal order 10*, simple relationships between two variables can be formed, leading to simple deductive logic and simple univariate tests of empirical truths. One understands by using deduction; respondents can logically test whether something meets the regulations. If experts know their legal role, they know their role is to testify in an unbiased fashion. But some prosecutors and attorneys want experts to testify in a biased fashion as we have shown previously (Dattilio, Commons, Adams, Gutheil, & Sadoff, 2006). The outcome of biasing is a variable and has a value. If an expert works for only one side of cases (e.g., only for plaintiffs), one outcome is that attorneys for this side will hire that person. That outcome has a high bias value. The predictive variables such as working for one side make the relationship formal. Working for one side is a value of the variable, which is working for one side versus working for both sides. The value of that variable predicts the perceived amount of bias. Hence two variables are related. This may be shown if it is found that the experience of testifying for one side is perceived to bias that expert. In signal detection theory, this bias is called the presentation effect (Bar-Ilanl, Keenoy, Levene, & Yaari, 2009). Prior history alters response bias in the same way as variations in signal probability or payoff. The roles are demarcated by regulations, where the regulations are stated or implied. This result allows for roles being clearly defined.

For example, consider endorsing the questionnaire item: "any prosecution-only criminal-case expert witness reveals bias". If the expert testifies in a biased way, then that testimony helps the lawyer win cases more often; if the lawyer wins more often, then the lawyer likes to hire those experts; experts who work for one side are hired by lawyers who like them that way.

At the *systematic order 11*, multiple relationships are coordinated. The role of the expert is to make judgments rather than following rigid rules and regulations.

At the *metasystematic order 12*, systems of relationships may be compared. One understands that an expert cannot be bias-free.

2. Method

2.1. Participants

Participants in this study were 46 attendees at the annual meeting of the American Academy of Psychiatry and Law (AAPL). They were 81.4% (35 out of 43) M.D.'s. There were 89.2% (33 out of 37) who were APA (psychiatry) members and 83.3% (30 out of 36) were members of AAPL. There were 9.3% (4 out of the 43) who were psychologists, 75% (3 out 4) of whom were a1so member of the APA (Psychological) and 1 of whom was in Division 41, American Psychology-Law Society. There were 2 Lawyers, 1 with an M.D. and 1 with a Ph.D. There was one person with a B.A. The average number of years in forensic practice was 11.34 (SD = 9.32) and the annual number of forensic cases was 48.82 (SD = 79.07). The study used a sample of convenience that is not likely representative of the general population of practicing forensic psychiatrists. The attendees were largely homogeneous. There was no reason to believe that the results were influenced by this small variation among the participants because of their high involvement in forensic matters.

2.2. Procedure

After obtaining permission through the Massachusetts Mental Health Center human studies committee and approval from the Research Committee of the AAPL, a questionnaire was handed out at the annual meeting of the American Academy of Psychiatry and Law (AAPL). Participants in this study were 46 attendees at one of the "professional issues in the twilight zone" workshops held at the

meeting. Participants voluntarily attended a workshop advertised as an opportunity to participate in both research and discussion of attorney–expert matters that were not often openly addressed, e.g., that existed in an insufficiently assessed "twilight zone". As some of those present had also attended similar workshops previously, and the basic theory and early results had been presented at the AAPL presidential address of 2000 (Gutheil, 2001), there was some familiarity with the format from at least some attendees.

After results were obtained, data analysis was performed. First, Rasch analysis was used to measure the perceived bias of items on a single one dimensional manner. Second, a factor analysis was performed. The factor analysis showed how the items are related to one another. Third, the inherent difficulty of the items was scored using the Model of Hierarchical Complexity Scoring Manual (Commons, Miller, Goodheart, & Danaher-Gilpin, 2007). This scoring was used to understand the result of the factor analysis as well as the degree of bias of items measured by Rasch analysis. Lastly, a regression analysis tested how much the Order of Hierarchical Complexity of the items predicted the degree of bias of each item measured by Rasch analysis.

2.3. Instrument

The instrument was constructed by the authors. Appendix A shows the instrument that was handed out to subjects at the above workshop. As seen in Appendix A, the questionnaire did not use the word "bias" in its prologue but instead identified "expert witness reactions to cases" as the focus of the study. Subjects were asked to think of recent cases in which they had served as expert witnesses as they answered the questions. The first series of queries addressed, on six point scales (mean 3.5), the issue of an expert's influence on case outcomes and the subjects' emotional reactions to those outcomes. Next was a series of queries that asked subjects to identify potentially biasing factors, from least biasing to most biasing, such as money, prestige, high profile cases, etc. The final series of queries focused on expert attitudes toward bias and biasing factors. The majority of questions were asked in regards to "opposing experts." This phrasing was used in order to minimize the desire to answer less than honestly about one's own actions if they were not seen as socially acceptable.

3. Results

3.1. Rasch analysis

Using a Rasch analysis, all but one item fit onto a single dimension of what the biasing potential was. To carry out the Rasch analysis (Bond & Fox, 2007), we put participants' answers to all of the questions pertaining to potentially biasing situations into a data file for analysis. The Rasch analysis retained 38 respondents. There were 8 non-respondents out of the original 46. They were dropped from further analysis because they provided no answers. For a Rasch analysis, this is a small sample. A sample size of 30 assures 95% confidence, and a sample size of 100 is recommended. The results were also analyzed to see if any outlier respondents needed to be dropped, but there were none. All the items but one clearly fit on a single perceived-seriousness-of-biasing scale (Linacre, 2002; Wright & Linacre, 2001). This scale and the placement of the various items are shown in Fig. 1. Note that more negative scores, which are shown at the bottom of Table 2, reflect a rating of greater biasing potential.

The item that was judged to produce the most extreme amount of bias in an expert was the frequency with which a respondent will turn down cases that evoke personal discomfort (Rasch score −1.55). Second to this was a group of situations in which it is known that certain expert witnesses testified consistently for only one side (e.g. plaintiff-only civil-case, −.89; prosecution-only criminal-case, −.87; defense-only criminal-case, −.83; defense-only civil-case, −.79). This finding

```
                    RATERS -MAP- QUESTIONS
                         <more>|<rare>
          2                    +   TREATER
                               |


                               |
                               |Q HIREDGUN
                               |

                               |
                               |
          1                    +   OWNLOSE
                               |   JUSTIFY
                               |S
                          4  Q|   ETHNIC    MENTAL     NOMENTAL
           1   7  13  32  40  |   OPHAPLOS  OWNWIN
              5  16  37  42 S|   LAWFIRM
        6  23  26  41  45  47  |   OUTCOME
    0  9  18  22  34  36  46 M+M GOALS     LAWYERID  MONEY     OPCOMLOS
          8  11  12  20  25  |   CASESIDE  DISTORT
             10  15  29 S|   SHOWOFF
          2  17  21  24  38  |   BIASFREE  IDEALS    OPHAPWIN  PROFILE
                    3  35  |   SURETY
                      27 Q|S OPCOMWIN
                               |   DEFCIVIL  DEFCRIME  PLAINCIV  PROCRIME
         -1                    +
                               |

                               |
                               |
                               |Q PERSONAL
                               |

                               |
         -2                    +
                         <less>|<frequ>
```
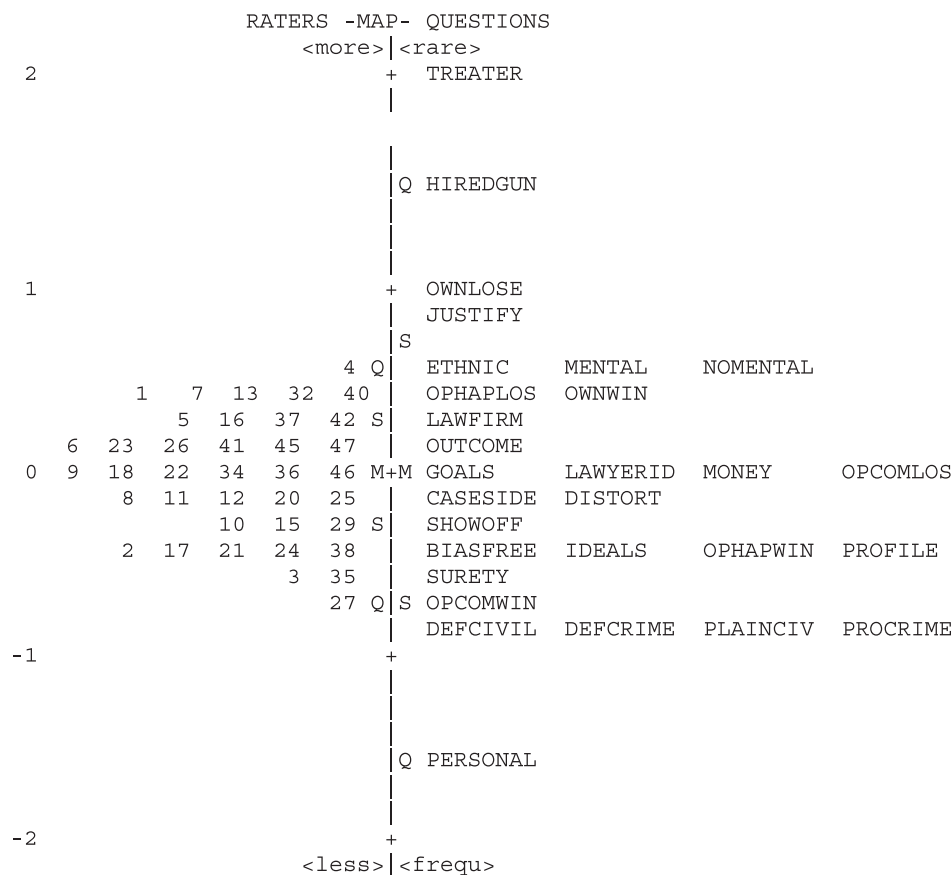
Fig. 1. Expert bias Rasch variable map. Input: 38 raters, 29 questions analyzed: 38 raters, 29 questions, 6 cats V2.90. In the Rasch map, the y axis represents the intensity of the perceived bias. The most biased items are at the bottom. On the right side are the item scores. On the left side are the person scores.

supports our contention that juries will most likely perceive expert witnesses as less biased if they have testified often enough for both sides of cases.

As stated above, each interval on a Rasch is of equal size; as a result equal differences in scaled measurements between biases represent equal differences in the resulting perceptions of severity. Thus, the perceived difference in bias between a respondent's assessment of the opposing expert witness as being a "hired gun" (1.6) and the compromise caused by an expert witness and parties on his/her side being of the same race or ethnic group (0.61) (1.6–0.61 = 1) is approximately the same as the perceived difference in bias between the respondent's assessment of the appropriateness of an opposing expert witness also being the examinee's treater (0.82) and the respondent's assessment of the biasing potential of an expert witness' persistent inclination to one side or another in court cases (−0.21) (0.81 − (−0.21) = 1). The difference between each set of biases is approximately one unit. That means that the two differences are of the same size.

A situation considered lowest in terms of biasing potential was, surprisingly enough, the one where the opposing expert had been the examinee's treater (1.99). Only slightly more biasing was the situation where the respondent had assessed the expert witness on the opposite side to be a hired gun (1.6). Also rated quite low was respondents' assessment of their own degree of happiness in a given case in which they had testified "appropriately," but that side of the case lost with a probably unjust outcome. "Possibly unjust outcome" as used in the instrument may itself be subject to interpretation bias (.97). Note that the scaled value of the difference between the opposition experts being the treater is (1.99) and the respondents' assessment of their own degree of happiness in a case that was lost (.97) was also about 1 unit.

There was one item that simply did not fit the scale. An item is considered not to fit when it had a very large "infit error" (Wright & Linacre, 2001). An infit error is a "fit" statistic that focuses on the size of the difference between the predicted performance of an item or person versus the observed value. A large infit error (2.0 or above) means that an item does not fit the scale well — there are large deviations of the observed rating from the expected ratings for that item (Linacre, 2002; Linacre, 2003). The item that did not fit the scale was one in which the respondent was asked whether they had ever decided to take action after concluding that an expert witness on the side opposite from the respondent's had acted unprofessionally during the course of a case. In constructing a single dimension scale, one draws from a large number of items. Some of these items help measure responses that fall on that scale. Some items will not "fit" the scale as described by Linacre (2002). In general, mean-squares infit errors greater than 2.0 degrade measurement of a Rasch scale; with this in mind, the item was not included. This item may not have fit because it lies on another dimension, one that reflects activity – perhaps professional disciplinary activity – outside of the court system.

3.2. Principal Components Analysis

A Principal Components Analysis (factor analysis) was performed. Items with coefficients of less than 0.45 were eliminated. Out of the original 24 items, 18 remained. Factor 1 accounted for 15.6% of variance and has 7 items with factor loading higher than 0.45. Factor 2 accounts for 12.0% of variance and has 7 items with factor loading higher than 0.45. Factor 3 accounted for 9.2% of variance and has 4 items with factor loadings of higher than 0.45. The three factors collectively explained 36.8% of total variance.

Table 2
Expert bias: 38 raters, 29 questions.

Item title	Questionnaire item	Scaled score	Infit Mnsq	Oder factor loading	Order
FOLLOWUP	Decided to take action after concluding that an expert witness on opposite side from respondent's has acted unprofessionally during the course of case	–	–		
TREATER	Been involved in a case where an expert witness on opposite side from respondent's has also been the examinee's treater	1.99	.52		
HIREDGUN	Assessed the expert witness on opposite side from respondent's in a case to be a "hired gun" (supplying an opinion only for money) during the course of a case	1.60	.31		
OWNLOSE	His/her own degree of "happiness" in a given case where, despite his/her having testified "appropriately," his/her side loses with the possibility of an unjust outcome	.97	.87	3	12
JUSTIFY	At the time of given events of the rectitude of opposing expert witness' also being the examinee's treater	.82	2.53	3	12
ETHNIC	How frequently an expert's objectivity is compromised when the opposing expert witness and his/her side are of the same race or ethnic group as the first expert witness	.61	.78		
NOMENTAL	Proportion of expert witnesses believing, to a degree that biases their testimony in insanity cases, crime almost never to be related to mental illness	.55	.67	2	11
MENTAL	Proportion of expert witnesses believing, to a degree that biases their testimony in insanity cases, all crime to be substantially related to mental illness	.52	.68		
OWNWIN	His/her own degree of "happiness" in a given case where, despite his/her having testified "appropriately," his/her side wins with the possibility of an unjust outcome	.50	.72	3	12
OPHAPLOS	Degree of "happiness" felt by opposing expert witness who has testified in an "appropriate manner" in a given case, if the latter's side loses	.41	.74		
LAWFIRM	Biasing potential of (retaining) law firm on the expert witness	.29	1.11		
OUTCOME	Agreement on whether favorability of outcome for his/her own side on a given case indicates the quality of respondent's work on that case	.18	1.05	1	10
LAWYERID	Biasing potential of an expert witness' identification with the (retaining) attorney	.06	.61	2	11
GOALS	Biasing potential of an expert witness' dedication to his/her social goals	.01	1.19	2	11
OPCOMLOS	Degree of competency felt by opposing expert witness who has testified in an "appropriate manner" in a given case, if the latter's side loses	.00	.87	3	12
MONEY	Biasing potential of money on expert witness	−.03	1.08		
DISTORT	Whether experts' claims of "compensating" for their own biases does not constitute striving for objectivity	−.12	1.14		
CASESIDE	Expert witness' inclination to one side (plaintiff/prosecution) or another (defense) in court cases	−.21	.92	2	11
SHOWOFF	Expert witness' desire to show off his/her own expertise, skill, erudition or the like	−.23	.79	1	10
OPHAPWIN	Degree of "happiness" felt by opposing expert witness who has testified in an "appropriate manner" in a given case, if the latter's side wins, instead	−.41	1.07		
BIASFREE	Frequency of opposing expert witnesses' belief in their own bias-freeness			−.41	.79	2
PROFILE	Biasing potential of the high profile of a given case on expert witness	−.43	1.05	1	9
IDEALS	Biasing potential of an expert witness' own "personal philosophy" on him/her	−.44	1.27	2	11
SURETY	Frequency of opposing expert witnesses' confidence in their own ability to compensate for obvious bias (e.g. for always working for one side in court cases)	−.56	.75	2	11
OPCOMWIN	Degree of competency felt by opposing expert witness who has testified in an "appropriate manner" in a given case, if the latter's side wins, instead	−.74	1.01		
DEFCIVIL	Agreement that any defense-only civil-case expert witness reveals bias	−.79	1.34	1	10
DEFCRIME	Agreement that any defense-only criminal-case expert witness reveals bias	−.83	1.47	1	10
PROCRIME	Agreement that any prosecution-only criminal-case expert witness reveals bias	−.87	1.23	1	10
PLAINCIV	Any plaintiff-only civil-case expert witness reveals bias	−.89	1.38	1	10
PERSONAL	Frequency at which respondent will turn down cases that evoke personal discomfort or squeamishness (due to personal attachments/involvements or memory of similar experiences)			−1.55	1.43
MEAN		.00	1.01		
SD		.76	.40		

Note. The scaled value is the Rasch scaled value. The more negative the value, the more serious the perceived bias. Infit errors are shown in the next column to the right. A probability expression can be calculated and used to combine any participant's estimated measure with any item's estimated measure to produce expected response values. These values can then be compared with the observed responses to detect misfitting responses and indicate potentially problematic items. Infit errors are the unweighted item and person fit statistics, which are sensitive to irregular inlying patterns. Outfit, shown in the next column, are the unweighted item and person fit statistics, which are sensitive to unexpected rare extremes.

These three factors were examined. Factor 1 items were concerned with the degree of bias due to high profile cases, working for only one side, judging the quality of respondents' work based on outcome and expert witness' desire to show off expertise. Factor 2 items were biased by personal philosophy, believing in one's own bias freeness, identifying with the attorney, social goals, holding a position that mental illness never causes insanity. Factor 3 represented items that were relatively unbiased. Results of the factor analysis are shown in Table 3.

3.3. Scoring for Orders of Hierarchical Complexity

On examination, the items within each factor shared a common characteristic called their Order of Hierarchical Complexity. The items were then scored to obtain that order. The way in which one goes about scoring is rather simple. First, one has to carefully describe what has to be done to complete the task successfully. Then one makes a guess as to the hierarchical complexity of that task based on the task analysis.

Then one checks that guess using the three axioms for a *higher Order of Hierarchical Complexity*. Please refer to Table 3 for a list of item orders.

3.3.1. Abstract order 9
The item scored as abstract order 9 is:

Respondent's assessment from professional experience of the biasing potential of the high profile of a given case on expert witness. This loads .638 on factor 1.

A high profile is an abstract order value of a variable. Responding to that shows sensitivity beginning at stage 9.

3.3.2. Formal order 10
All of the items at the formal order 10 are from factor 1. Items at the formal order 10 coordinate two variables. An example is that an expert working for only one side shows bias. As explained in the

Table 3
Factor analysis of items.

Component matrix[a]	Component		
	1	2	3
Respondent's level of agreement that any prosecution-only criminal-case expert witness reveals bias	0.863		
Respondent's level of agreement that any defense-only criminal-case expert witness reveals bias	0.818		
Respondent's level of agreement that any defense-only civil-case expert witness reveals bias	0.816		
Respondent's level of agreement that any plaintiff-only civil-case expert witness reveals bias	0.794		
Respondent's assessment from professional experience of the biasing potential of the high profile of a given case on expert witness	0.638		
Respondent's level of agreement on whether favorability of outcome for his/her own side in a given case indicates the quality of respondent's work on that case	0.519		
Respondent's assessment from professional experience of the biasing potential of an expert witness' desire to show off his own expertise, skill, erudition, or the like	0.511		
Respondent's assessment from professional experience of the biasing potential of an expert witness' own "personal philosophy" on him/her		0.772	
Respondent's assessment from professional experience of frequency of opposing expert witnesses' belief in their own bias-freeness		0.678	
Respondent's assessment from professional experience of the biasing potential of an expert witness' identification with the (retaining?) attorney		0.646	
Respondent's assessment from professional experience of the biasing potential of an expert witness' inclination to one side (plaintiff/prosecutorial) or another (defense) in court cases		0.606	
Respondent's assessment from professional experience of the biasing potential of an expert witness' dedication to his/her social goals		0.551	
Respondent's assessment from professional experience of frequency of opposing expert witnesses' confidence in their own ability to compensate for obvious bias (e.g., for always working for one side in court cases)		0.544	
Respondent's assessment of proportion of expert witnesses believing, to a degree that biases their testimony in insanity cases, crime almost never to be related to mental illness		0.449	
Respondent's assessment of degree of his/her own "happiness" in a given case where, despite his/her having testified "appropriately", his/her side wins with the possibility of an unjust outcome			0.653
Respondent's assessment at the time of the given events of the rectitude of opposing expert witness' also being the examinee's treater			0.558
Respondent's assessment of degree of competence felt by the opposing expert witness who has testified in an "appropriate manner" in a given case, if the latter's side loses			0.529
Respondent's assessment of degree of his/her own "happiness" in a given case where, despite his/her having testified "appropriately", his/her side loses with the possibility of an unjust outcome			0.451

Note. Extraction method: Principal Components Analysis.
[a] 3 components extracted.

Introduction, this item coordinates two variables: the behavior of working for one side, and the outcome of being hired more often. If the attorney wants the expert to work for one side, then the attorney will hire the expert with their bias. The items at the formal order are:

Respondent's level of agreement that any prosecution-only criminal-case expert witness reveals a bias. This loads 0.863.

Respondent's level of agreement that any defense-only criminal-case expert witness reveals a bias. This loads 0.818.

Respondent's level of agreement that any defense-only civil-case expert witness reveals a bias. This loads 0.816.

Respondent's level of agreement that any plaintiff-only civil-case expert witness reveals a bias. This loads 0.794.

Respondent's level of agreement on whether favorability of outcome for his/her own side in a given case indicates the quality of respondent's work on that case. This loads .519.

Respondent's assessment from professional experience of the biasing potential of an expert witness' desire to show off his own expertise, skill, erudition, or the like. This loads .511.

3.3.3. Systematic order 11

Items in factor 2 are at the systematic order 11. At the *systematic order 11*, multiple relationships at the formal order 10 are coordinated to form systems. The role of the expert is to make judgments rather than following rigid rules and regulations. Items that loaded highly on factor 2 were biased by personal philosophy, believing in their own bias freeness, identifying with the attorney, social goals, holding a position that there is never mental illness that causes insanity, etc. Five systems of relationships are described.

System 1: Personal philosophy and view
Personal philosophy and view may bias an expert witness. For example, some expert witnesses have to make judgments of whether the defendant is mentally ill. An expert witness may believe that a

person is legally insane when they cannot tell right from wrong, and when they do not have the capacity to know what they were doing. An alternative view would be that a person is insane when they could not conform their behaviors to the social norm and the law. These beliefs are at the systematic order, because personal philosophy is formed by formal order relationships. It is the product of analysis and abstraction of personal experiences, acquired knowledge and values. Items in System 1 are:

Respondent's assessment from professional experience of the biasing potential of an expert witness' own "personal philosophy" on him/her. This loads .772.

Respondent's assessment from professional experience of the biasing potential of an expert witness' dedication to his/her social goals. This loads .551.

Respondent's assessment of proportion of expert witnesses believing, to a degree that biases their testimony in insanity cases, crime almost never to be related to mental illness. Loads .449.

System 2. The experts believe themselves to be bias free, or the opposing experts have confidence in their own abilities to compensate for obvious bias (e.g., for always working for one side in court cases)

Judging the degree of bias of an opposing expert witness who believes in his or her bias-freeness is a task at the systematic order 11. One has to be able to know the cause of one's own bias, and one has to act on it. Then one feels that the bias is compensated, and one is bias-free. In order to do that, one has to evaluate one's own behaviors, or step outside of the self. As Rodriguez (1992) has shown, this is a task at systematic order 11. Items in this system are:

Respondent's assessment from professional experience of frequency of opposing expert witnesses' belief in their own bias-freeness. This loads .678.

Respondent's assessment from professional experience of frequency of opposing expert witnesses' confidence in their own ability to compensate for obvious bias (e.g., for always working for one side in court cases). This loads .544.

System 3. Identification with the lawyer

Tasks at this order coordinate two relationships. First, the expert witness identifies with the lawyer, and therefore takes the lawyer's side. Second, lawyer hires expert witnesses who identify with the lawyer. The item in this system is:

Respondent's assessment from professional experience of the biasing potential of an expert witness' identification with the (retaining) attorney. It loads .646.

System 4. Showing inclination to one side

Tasks at this order coordinate two relationships. First, expert witnesses who show inclination to one side of cases may help one side win more often. Second, winning more often causes expert witnesses to be hired more often. Item in this system is:

Respondent's assessment from professional experience of the biasing potential of an expert witness' inclination to one side (plaintiff/prosecutorial) or another (defense) in court cases. This loads .606.

3.3.4. Metasystematic order 12

Items in factor 3 are at the metasystematic order 12. At the *metasystematic order 12*, systems of relationships may be compared. Because the task of evaluating these items requires a very high stage, items on the factor 3 loaded less than the items on the first two factors. Their Rasch scores were also higher than the average, indicating that they are rated as less biasing than items in factors 1 and 2. Each item will be analyzed in the following:

Respondent's assessment of degree of his/her own "happiness" in a given case where, despite his/her having testified "appropriately", his/her side wins with the possibility of an unjust outcome. It loads .653.

Respondent's assessment of degree of his/her own "happiness" in a given case where, despite his/her having testified "appropriately", his/her side loses with the possibility of an unjust outcome. It loads .451.

These two items say that the person testifying is reinforced by the quality of that judgment and testimony and not controlled by the outcome of the case. To testify "appropriately" is likely to be a task at the metasystematic order, which coordinates legal principles, moral judgment and personal understanding of the case. That indicates the expert's behavior is principled and controlled by principled stage 12 reinforcers.

Respondent's assessment of degree of competence felt by the opposing expert witness who has testified in an "appropriate manner" in a given case, if the latter's side loses. It loads .529.

Again, this is about the quality of that judgment and testimony, not controlled by the outcome. That indicates the expert's behavior is principled and controlled by principled stage 12 reinforcers.

Respondent's assessment at the time of the given events of the rectitude of opposing expert witness' also being the examinee's treater. It loads .558.

This is a conflict of interest scenario. The expert witness' role is to give neutral opinions about the patient's state of mind, not to get involved in judging others. That is the job of the court. The examinee's treater, or therapist, should represent the interest of the patient. Judging the righteousness of the opposing expert witness is a task that requires understanding of the roles of the expert witness and the treater, both of which are systems at systematic order 11.

3.4. Regression analysis

The Rasch scores of the items and the Orders of Hierarchical complexity of the items tracked each other to some extent. A simple linear regression with the Order of Hierarchical Complexity as the independent variable and Rasch scores of the items as the dependent variable shows that $r(16) = 0.698$, $r^2 = 0.487$. Orders of Hierarchical Complexity significantly explains the variation in item Rasch scores, $F(1, 16) = 15.193$, $p = 0.001$.

4. Discussion

The participants' scaling of items reflects a rough logic and gives empirical support to a common sense view. First is the belief that the single most biasing situation is one in which the experts have some kind of a personal experience relating to the content of the case (enough so that they turn the case down). Second is that when experts work for only one side, they are much more likely to be perceived as biased. Third, situations such as one's personal philosophy, or whether a case is high profile, are not perceived as being comparably biasing.

The one possible exception to this common sense view is the low perceived biasing potential given to the litigant's treater serving also as expert. It has been argued elsewhere that the potential advocacy and subjectivity biases of the treater are among the strongest arguments for separating the expert and treater roles (Strasburger, Gutheil, & Brodsky, 1997). It is unclear why this widely known problem was not reflected on the scale here. It is possible that all other bias sources were regarded as more powerful; thus, treater-expert pairing was the "least worst" on this particular list. A possible explanation for the low perceived bias of a treater being an expert is that a treater is perceived to "know" the patient better than an expert who has spent much less time with the client. It may be that many of the participants condone or participate in the practice of serving as an expert in cases involving clients whom they see in treatment. It is also possible, but more problematic, that participants were not adequately aware of treater-expert biasing potential, a problem certainly shared by many attorneys. But this is unlikely given that the infit errors were not large, which suggests a uniformity of response to this item.

The highest scale item – turning down a case causing personal discomfort – would, indeed, be a classic and logical source of potential bias; ironically, this item takes care of itself, so to speak, since turning down such cases obviously removes the risk of biased testimony. Thus, an answer to the cross examination question, "Doctor, did not the fact that this malpractice case was brought against a fellow psychiatrist make you uncomfortable?" might be: "If it had made me noticeably uncomfortable, I would have turned down the case."

Time and again, the law requires expert testimony. Informally, judges and attorneys say that experts rarely have any effect on the outcome of cases. Having a scale of how biased an expert may be perceived allows us to show how serious a biasing factor may be, at least as perceived by experts. To deal with this on direct, attorneys might bring out that their witness has some biases, but they are considered weak. Then, on cross, one might be able to ask about the other side's expert and point out that other biases are large.

4.1. Are jurors' bias and experts' perception of bias similar?

Here we have focused on an expert's perception of bias regarding the opposing expert and suggested that this bias could be used in a critical cross-examination (Easton, 2000). There is some evidence that cross examination of defense experts on issues of bias rather than medical care may negatively affect an audience's perception of

the cross-examining attorney's case (Harrison, 2001). Also, juries are already biased against experts in general (Harrison, 2001; Shuman, Champagne, & Whittaker, 1996). A study similar in design to the present study, but with jury-eligible adults as respondents, is needed to assess directly the types of judgments "jurors" make about potential biasing situations. For those items that continue to show bias away from the center of the scale, a series of empirical studies might present controlled stimulus materials in hypothetical case vignettes to jury-eligible adults.

One of the most interesting outcomes of our study – which experts can supply to attorneys – is that a Rasch analysis shows when something does *not* belong on a scale, where another expert asserts that it does. The example in this study was taking action against an opposing expert for unprofessional conduct. On the face of it, one might think that this question would, indeed, fit on a perceived bias scale; but that was not our result. We can imagine that – in attempting to define the theoretical elements of the forensic field – locating comparable data on other scales might help separate empirical wheat (i.e., the relevant source of bias) from the chaff.

4.2. Factor analysis and stage

Here we show that the stage required to understand an item predicts how strong is the perceived bias of that item. This is probably what underlies our common-sense notions of bias. Somewhat to our surprise, the three factors obtained reflected the a priori difficulty of the items, which is still consistent with a one dimensional model because the factors just reflect the clumping of the items along that dimension. Factor analysis does support the view of the "clumps". Factoring shows that the predominant order is formal; next most frequent is systematic, etc. This is what the Model of Hierarchical Complexity would predict. The results support the theory that value and difficulty interact. These means that the perceived value, which biases choice, is partially determined by the Order of Hierarchical Complexity of the item, and therefore the difficulty in understanding the item. What was not done here was to have an independent manipulation of value, so that contribution of the value and difficulty could be gauged separately.

Bias is usually thought of as a dependent variable to be described and predicted. In the simplest sense, it represents a deviation from having a neutral value ascribed to a choice. In psychology and economics it is a choice-outcome-related dependent variable that alters the rating or probability of an action. Hence it belongs to the Behavioral Economics "value/reinforcement" paradigm. In behavioral economics, all bias reflects the probability of making a response (or the tendency to make a response) based on perceived value of the outcome. The more often one does something that has a positive outcome, the higher the value of doing it. This is true whether an expert is aware of this effect or not.

Acknowledgments

Appendix A. Study of expert reactions to cases

This study will ask you questions about the expert reactions to cases. The purpose of this study is to explore an empirical basis, to facilitate the discussion among members of our profession, and to further our understanding of forensic practice. Your participation is completely voluntary and your answers will remain anonymous.

Consent form

I understand the intent and requirements of this study. The information that I provide about my personal history, as well as the responses I give to these questions, will be strictly confidential. My identity will not be divulged in any discussion, lecture, address, or publication derived from this project. By responding to these questions, I agree to participate in this study.

_____Initial _____ Month _____ Day _____ Year

Thank you in advance for help and participation. If you are interested in knowing about the results of this study, please write your name, address and e-mail address on a separate sheet of paper, along with the phrase: "Would like to hear results of this study"

To begin, please check off or fill in the appropriate answers below.

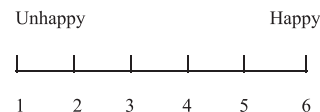|		
Gender	M	F
Preferred ethnic identifier _____		
Degree and discipline or specialization:		
M.D.	Specialization:_____	
Ph.D.	Discipline:_____	
M.S.W.	Discipline:_____	
J.D.		
Other (please specify): _____		
APA (Psychiatry) member	Yes____	No____
Board certified in general psychiatry:	Yes____	No____
Board certified in forensic psychiatry:	Yes____	No____
AAPL member:	Yes____	No____
Fellow:	Yes____	No____
Spouse is a member:	Yes____	No____
APA (Psychology) member	Yes____	No____
APA Division 41— The American Psychology-Law Society		
Member:	Yes____	No____
Board certified in forensic psychology:	Yes____	No____
Licensed in Clinical Psychology:	Yes____	No____
Other certifications	Yes____	No____
Specify:_____		
Years in forensic practice: _____		
Yearly, approximately how many forensic cases do you handle? _____		
Weekly, what percent of time, do you do:		
Forensic work? _____ Clinical work? _____		
Other work? (specify): _____		

We are studying expert witnesses' reactions and interpretations of case outcomes. Please do not identify yourself. Please try to answer *every* question by circling the number closest to your opinion. Thank you.

Think of some recent cases in which you have testified as an expert witness.

1) A case outcome favoring the side of the case that retained the expert is a measure of how good a job the expert did.

Disagree   Agree

1   2   3   4   5   6
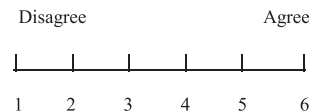
2) The case has been decided *in favor* of the side retaining the opposing expert who testified in an appropriate manner. To what degree does the opposing expert feel

Unhappy   Happy

1   2   3   4   5   6

2a) In the same situation, to what degree does the opposing expert feel

Incompetent   Competent

1   2   3   4   5   6
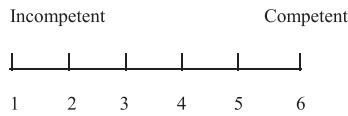
3) The case has been decided *against* the side retaining the opposing expert, who testified in an appropriate manner. To what degree does the opposing expert feel

Unhappy   Happy

1   2   3   4   5   6

3b) In the same situation, to what degree does the opposing expert feel

Incompetent   Competent

1   2   3   4   5   6
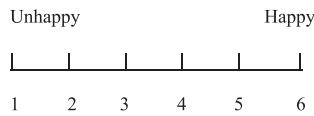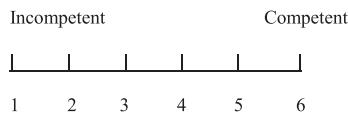
4) In a case in which you have testified appropriately, where the case was decided *against* the way you testified, but with a possibly unjust outcome, to what degree do you feel

Unhappy   Happy
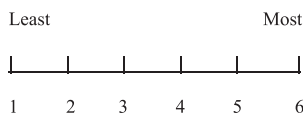
1   2   3   4   5   6
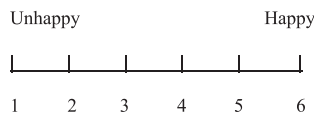
5) In a case in which you have testified appropriately, where the case was decided *in favor of* the way you testified, but with a possibly unjust outcome, to what degree do you feel
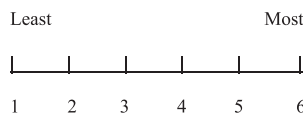
Unhappy   Happy

1   2   3   4   5   6
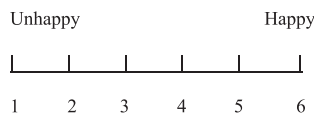
Based on your experience with opposing experts, *how potentially biasing was*:

6) Money

Least   Most

1   2   3   4   5   6

7) Prestige of law firm

Least   Most

1   2   3   4   5   6

8) High profile case

Least   Most

1   2   3   4   5   6

9) Experts "personal philosophy" (e.g., criminals are never insane, no one should get the death penalty)

Least   Most

1   2   3   4   5   6
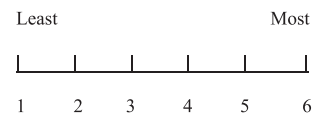
10) Side of the case requesting the expert

Least   Most

1   2   3   4   5   6

11) Identification with attorney

Least   Most

1   2   3   4   5   6

12) Social goal (e.g., stamp out sexual abuse) rather than goal of objectivity

Least   Most

1   2   3   4   5   6
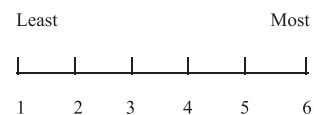
13) Demonstrate expertise, skill, erudition

Least   Most

1   2   3   4   5   6
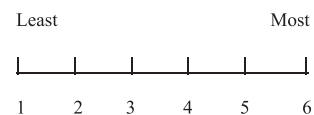
14) Someone, who always does criminal cases on the defense side, reveals a bias.

Disagree   Agree

1   2   3   4   5   6

15) Someone, who always does criminal cases on the prosecution side, reveals a bias.

Disagree   Agree

1   2   3   4   5   6
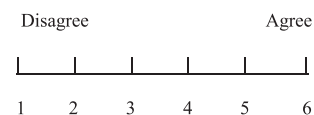
16) Someone, who always does civil cases on the plaintiff side, reveals a bias.

Disagree   Agree

1   2   3   4   5   6
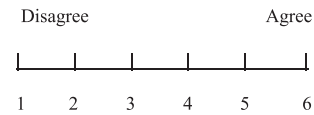
17) Someone, who always does civil cases on the defense side, reveals a bias.

```
        Disagree                    Agree
        |___|___|___|___|___|
        1   2   3   4   5   6
```
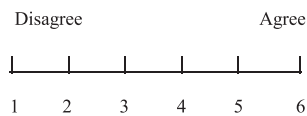
18) In insanity cases, what proportion of expert witnesses believe, *to a degree that biases their testimony*, that all crime is substantially related to mental illness?

```
         None                        All
        |___|___|___|___|___|
        1   2   3   4   5   6
```
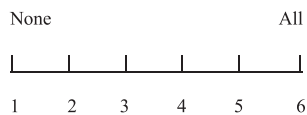
19) In insanity cases, what proportion of expert witnesses believe, *to a degree that biases their testimony*, that crime is almost never related to mental illness?

```
         None                        All
        |___|___|___|___|___|
        1   2   3   4   5   6
```
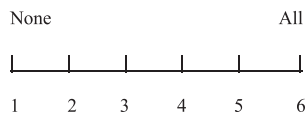
20) If you feel personal discomfort or squeamishness – the case is too close to home, the events involved you or someone close to you – do you turn down the case?

```
         Never                     Always
        |___|___|___|___|___|
        1   2   3   4   5   6
```
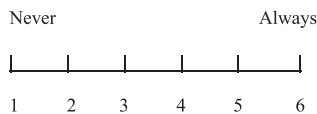
21) In a case where the opposing expert and that side's party are of the same race or ethnic group, is that expert's objectivity compromised?

```
         Never                     Always
        |___|___|___|___|___|
        1   2   3   4   5   6
```

22) In cases you have seen, does the other side's witness believe that he or she is bias free?

```
         Never                     Always
        |___|___|___|___|___|
        1   2   3   4   5   6
```

23) In cases you have seen, does the other side's witness believe he or she can compensate for an obvious bias (e.g. always working for the other side)?

```
         Never                     Always
        |___|___|___|___|___|
        1   2   3   4   5   6
```

24) When experts say they know they are biased but they compensate, they are merely "counter distorting," which does not constitute striving for objectivity.

```
        Disagree                    Agree
        |___|___|___|___|___|
        1   2   3   4   5   6
```

References

Bar-Ilanl, J., Keenoy, K., Levene, M., & Yaari, E. (2009). Presentation bias is significant in determining user preference for search results — A user study. *Journal of the American Society for Information Science and Technology*, *60*(1), 135–149.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). NJ: Lawrence Erlbaum.

Bradley, R. A., & Terry, M. E. (1952). Rank analysis of incomplete block designs. I. The method of paired comparisons. *Biometrika*, *39*(3/4), 324–345.

Commons, M. L. (1999). Threads of adult development. *Adult Development*, 1–2.

Commons, M. L. (2008). Introduction to the Model of Hierarchical Complexity and its relationship to postformal action. *World Futures: Journal of General Evolution*, *65*(1–3), 305–322.

Commons, M. L., & Miller, P. M. (1998). A quantitative behavior-analytic theory of development. *Mexican Journal of Experimental Analysis of Behavior*, *24*(2), 153–180.

Commons, M. L., Miller, P. M., Goodheart, E. A., & Danaher-Gilpin, D. (2007). Hierarchical Complexity Scoring System (HCSS): How to score anything. Retrieved from http://dareassociation.org/Papers/HCSS%20Applied%20to%20Terrorism.rtf

Commons, M. L., Miller, P. M., & Gutheil, T. G. (2004). Expert witness perceptions of bias in experts. *The Journal of the American Academy of Psychiatry and the Law*, *32*(1), 70–75.

Commons, M. L., Trudeau, E. J., Stein, S. A., Richards, F. A., & Krause, S. R. (1998). The existence of developmental stages as shown by the hierarchical complexity of tasks. *Developmental Review*, *18*, 237–278.

Dattilio, F. M., Commons, M. L., Adams, K. M., Gutheil, T. G., & Sadoff, R. L. (2006). Pilot Rasch scaling of lawyers' perceptions of expert bias. *The Journal of the American Academy of Psychiatry and the Law*, *34*(4), 482–491.

Easton, S. D. (2000). Ammunition for the shoot-out with the hired gun's hired gun: A proposal for full expert witness disclosure. *Arizona State Law Journal*, *32*, 465–614.

Gorsuch, R. L. (1983). *Factor analysis.* Hillsdale, NJ: Lawrence Erlbaum.

Gutheil, T. G. (2001). Adventures in the twilight zone: Empirical studies of the attorney–expert relationship. *Journal of the American Academe of Psychiatry and the Law*, *29*, 13–17.

Harrison, J. L. (2001). Reconceptualizing the expert witness: Social costs, current controls and proposes responses. *Yale Journal on Regulation*, *18*, 254.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases* (1st ed.). New York, NY: Cambridge University Press.

Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions*, *16*(2), 878.

Linacre, J. M. (2003). *Winsteps: Multiple-choice, rating scale and partial credit Rasch analysis.* Chicago, IL: Mesa Press.

Luce, D. (1959). *Individual choice behaviors: A theoretical analysis.* New York: J. Wiley.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests.* Chicago: University of Chicago Press.

Rodriguez, J. A., (1992). *The adult stages of social perspective-taking: Assessment with the doctor–patient problem.* (Unpublished Doctoral Dissertation). Harvard Graduation School of Education, Cambridge, MA.

Shuman, D. W., Champagne, A., & Whittaker, A. (1996). Juror assessments of the believability of expert witnesses: A literature review. *Jurimetrics Journal*, *36*, 383.

Strasburger, L. H., Gutheil, T. G., & Brodsky, A. (1997). On wearing two hats: Role conflict in serving as both psychotherapist and expert witness. *The American Journal of Psychiatry*, *154*(4), 448–456.

Wright, B. D., & Linacre, J. M. (2001). Glossary of Rasch measurement terminology. *Rasch Measurement Transactions*, *15*(2), 824–825.