

Arthur MacNeill Horton, Jr.
Cecil R. Reynolds *Editors*

Detection of Malingering during Head Injury Litigation

Third Edition

 Springer

Detection of Malingering during Head Injury Litigation

Arthur MacNeill Horton, Jr. • Cecil R. Reynolds
Editors

Detection of Malingering during Head Injury Litigation

Third Edition

 Springer

Editors

Arthur MacNeill Horton, Jr.
Neuropsychology Section
Psych Associates of Maryland
Towson and Columbia
Bethesda, MD, USA

Cecil R. Reynolds
Department of Educational Psychology
Texas A&M University College Station
Austin, TX, USA

ISBN 978-3-030-54655-7

ISBN 978-3-030-54656-4 (eBook)

<https://doi.org/10.1007/978-3-030-54656-4>

© Springer Nature Switzerland AG 2021

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The issue of potential malingering (or poor effort) in the context of head injury litigation has seen an explosion of research and commentary in the last 40 years. Various journals devoted to clinical neuropsychology practice of the 1950s, 1960s, and 1970s rarely published papers focused on the issue of malingering in civil litigation. The few papers that were published deal with malingering in regard to military service and criminal prosecution. It might be recalled that clinical neuropsychology began as an academic research endeavor.

Only after empirical studies proved the validity of neuropsychological evaluations as measures of brain behavior relationships was the necessary basis of scientific findings to support clinical applications established. Ralph Reitan, Oscar Parsons, Arthur Benton, and a bevy of their former students (Lawrence Hartlage and Charles Mathews, among others) published empirical research studies demonstrating the importance of clinical neuropsychology research to patient care. Their work formed the basis for clinical practice in neuropsychology. An interesting clinical note is that Dr. Benton initially developed his famous Visual Retention Test based on his clinical work during World War II, when Dr. Benton was based at the San Diego Veterans Administration Hospital and had to assess possible malingering by sailors who did not wish to return to fighting in the Pacific.

This growth of clinical neuropsychology research, clinical practitioners, and concurrent founding of journals to support scholarly inquiry and clinical practice has produced many revelations about the effects of closed head injury, an injury which was for many years believed to be of little consequence and one known as a silent epidemic. Coupled with the explosive growth of knowledge about the potential deleterious effects of closed head injury there has been increased personal injury litigation, changes in protocols for screening, assessment, and treatment of sports-related and war-related head injuries, and product liability suits. As more and more head injuries have come to be litigated and the potential sums of money involved have become enormous, issues and concerns about malingering (or poor effort) have grown substantially. By the 1980s, detection of malingering and its evaluation had found a routine place in the primary journals of neuropsychology.

The first edition of this book attempted to address the problem of malingering in head injury litigation. Several things were clear at the time of the first edition. The first was that malingering was a very substantial problem in head injury litigation. For example, empirical research findings had indicated that at least 25% of cases of head injury in litigation involve malingering. Second, the possibility of malingering existed in all head injury litigation cases and clinicians needed to be prepared to make the diagnosis when it was the most probable diagnosis. Third, there were many methodological, conceptual, and logistical caveats related to the detection of malingering. Fourth, there are emotional difficulties in labeling a patient a malingerer.

Malingering is a diagnosis with clear negative implications and is tantamount to calling a potentially brain injured patient a liar, something that can have very negative personal and financial consequences. Because of the very serious consequences, very convincing evidence is required for such a diagnosis and more than is the case for many clinical diagnoses made on a routine basis. The first edition attempted to demonstrate the utility and the pitfalls of various actuarial and clinical approaches to the diagnosis of malingering and equip the clinician with the necessary tools, knowledge, and logic to consider malingering and its alternative diagnoses intelligently, honestly, and ethically.

In the years that have passed since the first edition was published, much had happened in terms of research and clinical practice related to the detection of malingering in head injury litigation. For example, a new common practice was to use preferentially the term “poor effort” rather than “malingering,” as poor effort is a behavior that can be observed objectively. Research has shown in adults that effort explained 50% of the variance in the whole neuropsychological test battery but years of education explained only 12%, and severity of brain injury and age explained 4% and 3% respectively (Green, Rohling, Lees-Haley, & Allen, 2001). Also in children Kirkwood, Yeates, Randolph, and Kirk (2011) found while failure rates on effort tests were lower in children than in adults, effort testing explained 38% of the variance in the neuropsychological test battery.

On the other hand, the term malingering has been thought to require the forming of a conscious intention which as yet is unobservable in addition to a behavior. While it is clear that in the future it may be possible to determine if an individual has formed an intention to malingering as of the time, it has not been scientifically established. With the above concern noted, the term malingering is used in the title of this volume but the editors are aware of its limitations.

In the second edition of this work, the assembled chapters were based on rigorous scientific research but were also clinically oriented to facilitate their application to practice. Opening chapters disclosed the methodological and conceptual problems in the diagnosis of malingering to establish a clear mind-set of critical analysis before reading about methods proposed by other authors. The chapters that followed provided then current methods and thinking on multiple approaches to the detection of malingering during head injury litigation, including specific symptoms such as memory loss to more global claims of diffuse loss of function to cognitive and psychomotor arenas. The various presentations ranged from the strong actuarial

methods to careful, consummate clinical reasoning. The second edition, similar to the first edition, had been developed for the thoughtful, serious clinician who may be involved in evaluating patients with head injury who often become involved in litigation with regard to these injuries.

In the years since the second edition was published, there has been considerable additional research regarding malingering. A major conceptual change has been the distinction between performance validity and symptom validity. Briefly performance validity refers to malingering or effort tests that assess physical performance, be it motor or verbal. For example, the Word Memory Test (WMT) or Test of Malingered Memory (TOMM) would be examples of performance validity tests. In contrast, symptom validity refers to self-report that may be false for the purpose of feigning. The classic example of a symptom validity test is the various validity scales on the Minnesota Multiphasic Personality Inventory (MMPI) (in various editions such as the MMPI-2, MMPI-2-RF, and MMPI-3).

In addition, recent years have shown a great explosion in tests of effort using patterns of performance and cutting scores on neuropsychological tests not initially designed to assess effort to assess effort. Moreover, the large volume of clinical research has prompted many strong ideas and creative approaches and new methodologies to the detection of malingering. Due to the efforts of multiple researchers the diagnosis of malingering has profited from the establishment of a greater empirical scientific basis for decision-making.

The diagnosis of malingering is still, however, fraught with conceptual, philosophical, and logistical potholes. Because much has happened in the research basis for this clinical and forensic area, it is felt that a new third edition is required to address new research findings and changes in clinical practice that have occurred since the publishing of the second edition of this book.

This new third edition is intended to address new research findings and changes in clinical practice that have occurred since the publishing of the second edition of this book and provide practitioners with the necessary contemporary scientific findings to guide their clinical work and ensure that their patients receive the highest quality of clinical neuropsychological services.

Because of the increase in new research, it was decided to enlarge the third edition into two volumes. In the first volume, the authors (Faust, Gaudet, Ahern, and Bridges) discuss the complex methodological and conceptual problems in the diagnosis of malingering to clearly establish a mind-set of critical analysis before reading about issues and methods proposed by other authors. In the second volume, the authors address ethical issues (Kaufman and Bush), cultural aspects (Braw), and neuroimaging (Bigler). Specific test focused research is provided in chapters that follow related to the Word Memory Test (WMT), Medical Symptom Validity Test (MSVT) and Nonverbal-Medical Symptom Validity Test (NV-MSVT) (Armistead-Jehle, Denney, and Shura), Test of Malingered Memory (TOMM) (Perna), executive functioning tests (Suhr, Bryant, and Cook), and the MMPI-2, MMPI-2-RF, and MMPI-3 (Tylicki, Tarescavage, and Wygant). The next chapter is focused on assessing malingering in pediatric evaluations (Clegg, Lynch, Mian, and McCaffrey). The

last chapter focuses on methods and techniques for applying the information in the earlier chapters in forensic litigation (McCaffrey, Mian, Clegg, and Lynch).

The editors must express their appreciation to the chapter authors, who have made important contributions to the evaluation of malingering (or poor effort). Each has provided original insights, methods, and commentary on these very complex and difficult issues. Their willingness to share in the movement toward advancement in the diagnosis of malingering is greatly appreciated. To our editor at Springer, we would like to express our appreciation for the continuing faith in our efforts to produce a work that contributes significantly to the growth of clinical neuropsychology and the appropriate diagnosis of malingering (or poor effort). To the Springer staff and production editor, we also thank you for bringing the manuscript to its published conclusion with such promptness. To our two long-suffering wives, Mary W. Horton and Dr. Julia A. Hickman, goes our continuing and unfaltering love and appreciation for their help, support, kindness, and understanding during those times devoted to manuscripts such as this that pull from time otherwise spent together. We love you and thank you; we thank you very much!

Bethesda, MD, USA
Austin, TX, USA

Arthur MacNeill Horton
Cecil R. Reynolds

Conflicts of Interest Dr. Horton was a test consultant for the development of the Pediatric Performance Validity Test Suite (PdPVTS) and revision of the Test of Memory Malingering (TOMM) and received honoraria for his service but does not receive income from sales of the PdPVTS or the revised TOMM.

Dr. Reynolds is an author of the PdPVTS along with numerous other psychological and neuropsychological tests and receives royalties from their sales.

References

- Green, P., Rohling, M. L., Lees-Haley, P. R., & Allen, L. M. (2001). Effort has a greater effect on test scores than severe brain injury in compensation claimants. *Brain Injury, 15*(12), 1045–1060.
- Kirkwood, M. W., Yeates, K. O., Randolph, C., & Kirk, J. W. (2011). The implications of symptom validity test failure for ability-based test performance in a pediatric sample. *Psychological Assessment, 24*(1), 36–45. <https://doi.org/10.1037/a0024628>.

Contents

Assessment of Malingering and Falsification: Continuing to Push the Boundaries of Knowledge in Research and Clinical Practice	1
David F. Faust, Charles E. Gaudet, David C. Ahern, and Ana J. Bridges	
Ethical Issues in Assessing Malingering in Head Injury Litigation	157
Noah K. Kaufman and Shane S. Bush	
Cultural Aspects in Assessing Malingering Detection	177
Yoram Braw	
Neuroimaging and Invalid Neuropsychological Test Performance	201
Erin D. Bigler	
Use of the Word Memory Test (WMT), Medical Symptom Validity Test (MSVT) and Nonverbal Medical Symptom Validity Test (NV-MSVT) in Assessment Following Head Injury	223
Patrick Armistead-Jehle, Robert L. Denney, and Robert D. Shura	
Nonverbal Performance Validity Testing: Test of Memory Malingering (TOMM).	245
Robert Perna	
Malingering of Executive Functioning in Head Injury Litigation	257
Julie A. Suhr, Andrew Bryant, and Carrie Cook	
Assessment of Malingering Among Head Injury Litigants with the MMPI-2, MMPI-2-RF, and MMPI-3	289
Jessica L. Tylicki, Anthony M. Tarescavage, and Dustin B. Wygant	

Performance Validity Assessment in Pediatric Evaluations 309
Rachel A. Clegg, Julie K. Lynch, Maha N. Mian,
and Robert J. McCaffrey

**Explaining Performance and Symptom Validity Testing
to the Trier of Fact** 337
Robert J. McCaffrey, Maha N. Mian, Rachel A. Clegg,
and Julie K. Lynch

Index 353

Assessment of Malingering and Falsification: Continuing to Push the Boundaries of Knowledge in Research and Clinical Practice



David F. Faust, Charles E. Gaudet, David C. Ahern, and Ana J. Bridges

How can one make both a false-negative and a valid-positive identification simultaneously? Co-occurring correct and incorrect judgments can result either by identifying an injured individual who is also exaggerating deficit simply as a malingerer, or by identifying that same individual only as injured. In the first instance one misses the injury while correctly identifying malingering, and in the second instance one correctly identifies the injury but misses malingering.

As this example illustrates, the assessment of falsification or malingering often does not fall into neat packages. Impressive advances have led to the development of better methods, better strategies, broader options, enhanced awareness, and greater understanding, with psychologists and neuropsychologists easily being the most productive contributors to these noteworthy developments. However, critical problems and diagnostic puzzles remain, and as is often true as science advances, those problems tend to be considerably deeper and more complex than might first be realized. There is still a great deal more to learn about this domain, and in this volume we try to contribute in some small way to this endeavor. Ultimately,

D. F. Faust (✉)

Department of Psychology, University of Rhode Island and Department of Psychiatry and Human Behavior, Alpert Medical School of Brown University, Kingston, RI, USA
e-mail: faust@uri.edu

C. E. Gaudet

Department of Psychology, University of Rhode Island, Kingston, RI, USA

Psychology Service, VA Boston Healthcare System, Boston, MA, USA

Department of Psychiatry, Harvard Medical School, Boston, MA, USA

D. C. Ahern

Department of Psychiatry & Human Behavior, Alpert Medical School of Brown University, Providence, RI, USA

A. J. Bridges

Department of Psychological Science, University of Arkansas, Fayetteville, AR, USA

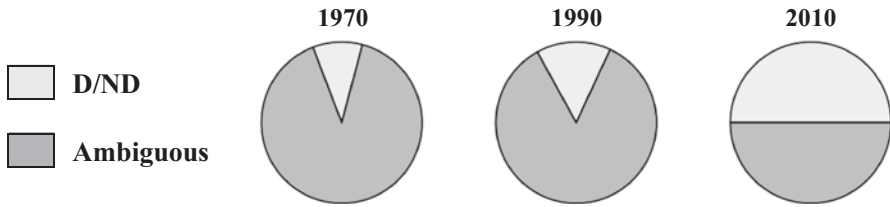


Fig. 1 Progress in increasing the proportion of Definitive or Near-Definitive (D/ND) cases

improved understanding and methods serve equally to identify false claims and verify true ones, and thus enhance the capacity of our profession to assist in such important tasks as the just resolution of legal conflicts, which is the normative role for expert witnesses.

One way to represent scientific progress is to divide pertinent cases into those that can be identified with certainty or near certainty versus those that remain ambiguous or difficult to identify and to look at changes in the proportions of these categories over time. We will refer to the former type of case as D/ND (definitive or near-definitive) and the latter as AMB (ambiguous). Of course, we are dichotomizing matters that lie on a continuum, but for current purposes finer divisions or more precise boundaries are not required because the intent is mainly conceptual. As shown in Fig. 1, suppose we traced the distribution of cases over the last 4 decades as follows, while presuming the level of ambiguous cases continues to gradually decline as of today.

We do not wish to debate the specific divisions across the pie charts for the moment. Given the accuracy rates that many studies yield, a reader might reject the proportions in the pie charts as misleadingly low, especially in the chart for 2010. We are not claiming that the proportions should be taken literally, the intent here being to illustrate progress over time. With that said, for reasons we will later address extensively, the results of many research studies, although certainly positive and encouraging, may substantially overestimate accuracy rates. In particular, many such studies primarily involve relatively clear or extreme cases as opposed to more ambiguous or difficult cases. Whatever one's position on these matters, we believe there would be broad consensus about the positive trends represented in the successive charts and the expectation that further gains have been made post-2010 and continuing to the present time.

As scientific knowledge has advanced, the percentage of cases that can be identified with high levels of accuracy has increased, with particular acceleration in progress during the last few decades as the level and quality of research have shown remarkable growth. The more we can whittle away at the remaining ambiguous cases (whatever their estimated frequency might be), the better off we will be, and it is sensible to focus research efforts on the types of cases that, despite our efforts so far, remain ambiguous or difficult. We might anticipate that these sorts of cases can present considerable scientific challenges, for if they were easy we would already know how to identify them. In many domains (e.g., golf, budget cutting, work efficiency), further advances can become progressively more difficult for various reasons, in particular because one can start with components that are easier to correct and because initial low levels of proficiency leave greater room and opportunity for gain. Without

losing sight of the impressive strides that have been made, the main focus of this volume is on these remaining ambiguous cases, not because we wish to concentrate on the negative but because they are a key to advancing proficiency—to achieving positive gains. Such cases often create significant scientific challenges and will require concentrated effort at least comparable to that which has already been expended. However, we think the prospects for further advance are good and that the effort is well justified given the importance of the problem.

Two areas of focus are critical to advance, and discussing them briefly at this juncture should provide a flavor for the sorts of matters we will cover. One is increased study of an underrepresented yet common group in litigation—those who are brain injured *and* falsifying. (Researchers studying psychological disorders have been giving more attention to co-presenting conditions for a number of years now, despite the challenges involved, and we believe it would be wise to do so for co-occurrences or co-phenomena in the area of falsification and malingering as well.) Unless one takes the extremist view that any and all falsification renders a person undeserving of any compensation (i.e., that the deserved retribution or consequence is the complete negation of any meritorious claims), a position we believe holds individuals to a standard of near-infallibility or moral perfection, then this group deserves our attention. Whatever our personal views on the matter, the outcome that should result when there is both legitimate injury and falsification has occupied and will occupy the trier of fact daily in courtrooms across the country, and it is an area in which mental health professionals could play a very important role in fostering more informed decisions, if and when sufficient research progress is made.

Second, our seemingly bright prospects for scientific advance in the appraisal of falsification hinges to no small extent on recognizing and correcting what we call the *extreme group problem* in research. Much contemporary research may not go far in reducing the percentage of ambiguous cases and may even produce the opposite result (i.e., lead us to miss cases we might identify correctly otherwise). These negative consequences stem largely from sampling problems in research, which result in groups that differ quantitatively and qualitatively from the remaining ambiguous cases. As we will argue, the extreme group problem is a common, highly impactful, yet often subtle methodological flaw. It is especially pernicious because the extent of the flaw may often be the most powerful influence on the accuracy rates obtained in studies, that is, the worse the flaw, the better a method seems to perform. When there is a powerful (or predominant) *positive* association between the magnitude of a design flaw and obtained accuracy rates, and this flaw goes unrecognized, a multitude of serious negative consequences are likely to follow. We will describe how the extreme group problem can be parsed and possibly corrected, although it may require substantial conceptual reframing, new avenues of research, and new metrics to detect, measure, and attenuate or negate its effects.

Our aim is not to critique the now considerable body of literature study-by-study, nor to address fundamental methodological points that have been cogently and convincingly described in the literature. Rather, our main intent is conceptual and prospective, with a particular focus on critical problems that may be under-recognized and suggestions and strategies that may assist in tackling challenging methodological hurdles.

1 Limitations of Experience in Learning to Detect Malingering: Benefits of Augmenting Clinical Judgment with Formal Methods

The intensity of reaction sometimes seen when research has raised questions about clinicians' capacity to detect malingering, especially absent the use of specialized methods and when depending primarily on subjective or professional judgment, seems to have quieted down as mounting scientific studies have made matters increasingly clear. Even more than 20 years ago, based on the additional evidence collected by that time, Williams (1998) put the matter thusly:

The study of malingering has moved beyond the controversies about whether clinicians are able and willing to detect it... the developing literature clearly suggests that clinicians using conventional strategies of interpretation cannot detect malingering and need some new systematic approach to the interpretation of conventional tests or new specialized symptom validity tests. (p. 126)

Although one might have preferred a different descriptor than "cannot detect malingering" such as "may have considerable difficulty" or "are highly prone to error," the same basic conclusions are echoed in more tempered form in the National Academy of Neuropsychology's position paper on malingering detection (Bush et al., 2005) and the American Academy of Clinical Neuropsychology's publication on this same topic (Heilbronner, Sweet, Morgan, Larrabee, & Millis, 2009). In these sources one will find statements such as "[U]se of psychometric indicators is the most valid approach to identifying neuropsychological response validity" (Heilbronner et al., 2009, p. 1106) and "[S]ubjective indicators, such as examinee statements and examiner observations, should be afforded less weight due to the lack of scientific evidence supporting their validity" (Bush et al., 2005, p. 424). Research supporting such statements includes studies demonstrating the difficulty of detecting lies or misrepresentations, the limits of experience and clinical judgment in learning to detect and identify malingering, and the potential and sometimes sizeable benefits realized when specialized methods are applied meticulously and interpreted in strict accord with scientifically based, formal decision procedures (see Faust, 2011, Chaps. 8 and 17).

Nevertheless, experience often has a powerful pull on clinical judgment and decision making. Given the inflated impression of efficacy that can easily result from experientially based impressions and its potential *detrimental* effects on accuracy in malingering detection when it overrides the use of more effective methods, the limitations of learning via experience in this domain are worth examining. One can start by considering the conditions that promote or inhibit experiential learning (Dawes, 1989; Faust, 1989; Faust & Faust, 2011). Experiential learning tends to be most successful when feedback is immediate, clear, and deterministic. By *deterministic*, we mean that the feedback is unfailingly or perfectly related to its antecedent, in particular the accuracy of judgments or conclusions. Thus, each time we are right we find out we are right, and each time we are wrong we are informed so. At the other end of the spectrum, learning can be difficult or impossible when no feedback

is received. In between, as the error term in feedback increases, that is, as the level of noise and inaccuracy in feedback grows, the more difficult learning tends to become.

The Category Test (Reitan & Wolfson, 1993) can serve to illustrate these points. Following the examinee's response, immediate feedback informs the person in no uncertain terms whether the response is correct. The feedback is deterministic: each time a response is correct a bell rings, and each time it is wrong a buzzer sounds. These are excellent conditions for learning from experience, and most examinees benefit greatly from the feedback, performing well above chance level. Further, if normal individuals were given the chance to take the Category Test again and again within a brief period of time, many would rapidly move toward very high levels of accuracy.

Imagine, however, a situation in which feedback is often no longer an easily distinguished bell or buzzer but something that perhaps sounds a little more like a bell than a buzzer or a little more like a buzzer than a bell. Imagine further that in many instances feedback is delayed, perhaps by minutes or hours or days, and that in the interim intervening events might occur that could alter the seemingly simple association between response accuracy and feedback. For example, in some instances some distorting influence might occur which leads a response of 2 to be misrepresented as 3, with feedback given accordingly. Imagine if, in addition, the feedback is systematically skewed in some fashion; for example, if the examinee is repeatedly informed that a certain type of misconception is instead correct. Imagine further that at times, perhaps more often than not, no feedback is given at all. Obviously learning via experience would become much more difficult, and one might welcome a community of scientists mounting a concentrated effort to unlock the keys to the Category Test.

We do not think it is overstating things to say that a clinician who depended solely on experience to learn malingering detection would be faced with much the same conditions as someone trying to learn under conditions of sporadic, skewed, delayed, noisy, and all too often misleading feedback. In many, if not most, instances, the clinician does not receive feedback on the accuracy of positive or negative identifications of malingering. When feedback is obtained it is often delayed, ambiguous, and skewed or distorted. If the clinician falsely diagnoses brain dysfunction, it would be the rare event for someone who is malingering to correct the misimpression. If the clinician falsely diagnoses malingering, then a plaintiff's sincere claims of disorder have not been believed in the first place, and subsequent sincere disagreement, should the plaintiff learn of the clinician's conclusion and have a chance to dispute it, are likely to be similarly rejected. The outcome of a courtroom trial, should the case be one of the small percentage that ever get that far, does not necessarily indicate the true answer, and can be contaminated by the clinician's own input. Although a clinician who believed the claimant was sincere might be confronted at trial with a videotape that provides convincing evidence that the practitioner was fooled, it establishes little other than judgmental fallibility rather than perfection, something that all but the most foolishly arrogant already recognize.

The attempt to identify and apply malingering indicators via experience, or perhaps to modify formally validated procedures on this same basis, encounters major obstacles. If one does not consistently know who are and are not the malingerers among those one evaluates, how can one determine the relative frequency of potential indicators across the target and nontarget groups? Even if such identifications are possible in some cases, absent a representative sample of cases, as opposed to the sample and distribution of cases the clinician happens to see in his or her setting, differential frequencies may be substantially misrepresented. An accurate appraisal of these differential frequencies is necessary to determine whether a sign is useful, just how useful it might be, how it compares with other signs, whether it should be included with other available predictors, and how it is to be combined with other predictors. As the Chapmans' original research (Chapman & Chapman, 1967, 1969) and much work thereafter has shown (Nickerson, 2004; Wedding & Faust, 1989), it can be very difficult to determine the association between variables, such as potential signs and disorder, in the course of clinical practice and observation. We are prone to forming false associations between signs and disorder and overestimating the strength of associations.

If and when valid signs are identified, one then wishes to adjust, as needed, the manner in which they are used or the cutting scores that are applied in accord with the relative frequencies of the target and nontarget populations in the setting of utilization. A decision rule that is effective in a setting with a very high rate of malingering will probably lead to far too many false-positive identifications if applied unchanged within a setting with a much lower frequency. As we will take up in greater detail later, decision rules should be adjusted in accord with frequencies or base rates in the setting of application (Meehl & Rosen, 1955). Optimum cutting points shift depending on the frequency of conditions.

The task that faces the clinician who tries to learn malingering detection via experience is thus as follows: The clinician needs a way to determine true status, determine the differential frequency of the target and relevant nontarget groups in the setting of interest, obtain representative samples of these groups, separate the valid and invalid signs through adequate appraisal in these groups, and then devise a proper means for combining the range of valid predictors that have been uncovered, preferably by considering such matters as their nonredundant contribution to predictive accuracy and the extent to which predictions should be regressed. To say the least, this is a formidable task. It is also one that creates a blueprint for researchers.

Some readers have undoubtedly pondered the various parallel problems that researchers routinely encounter in studies on malingering. For example, in many studies one cannot determine the true status of group members with even near certainty (e.g., whether those in the "malingering group" are really malingering). The same conditions required for learning through clinical experience need to be met for learning through research, and to the extent that studies fall short, the pragmatic help they can provide to clinicians will be compromised. Of course, this does not justify the stance that, because such conditions are imperfectly met by one or another investigation, one can then resort to experiential learning in which one routinely compounds, to a far greater extent, the methodological shortcomings of

research studies. We will address various problems that researchers face at length below, but would note here that the parallels are not complete. As is well known, researchers have a range of methods that may neutralize, attenuate, or gradually lessen impediments to learning or the enhancement of knowledge (e.g., greater opportunities to gather appropriate samples, use of control groups, implementation of various procedures to attenuate bias, opportunities to alter variables systematically, and greater luxury of trial and error learning).

2 Potential Benefits of Experience and Case Study

The preceding statements should not be confused with the view that clinical experience and impressions are of no use. Rather, it is important to recognize the strengths and limitations of such evidence. Perhaps the foremost concern with case study and related methods is one of sampling. As we will argue, sampling problems often also plague other research methods for investigating malingering, but they are especially acute with case study methods and typically render attempts at generalization on this basis alone as unwise, if not unwarranted and potentially irresponsible. Despite this critical limitation, it is also the case that clinical observation has led to brilliant insights, and it is sometimes hard to imagine how such ideas could have evolved in any other context. It seems almost pedantic to say that all forms of evidence do not serve all masters equally well. When evaluating malingering research, we need not apply criteria rigidly across a diverse set of contexts where they are not fully or at all appropriate. A related error would be assuming information that meets evaluative criteria in one context will do so across other contexts without considering the shift in epistemic standards that may be necessitated by context and intended use.

Although the distinction is somewhat artificial and the boundaries not always clear-cut, it is still helpful to distinguish what Reichenbach (1938) referred to as the context of discovery and the context of justification. To detect malingering, the clinician needs efficacious predictors. Of course, predictors that no one has ever thought of cannot be validated or applied. Surely no philosopher of science would suggest that the researcher “only identify potential predictors that are known in advance to be highly valid”; we are aware of no method for doing so and such a prescription would impossibly hinder investigation. More reasonable epistemic advice might be something like, “Test your best ideas or conjectures about potential predictors, and try to avoid potential predictors that have very little chance of success, unless you are totally impeded, or unless improbable indicators, should they pan out, are likely to be very powerful; but don’t inhibit yourself too much because it’s hard to anticipate nature and occasionally a seemingly outlandish idea turns out to be highly progressive.” In the context of discovery, one exercises considerably greater leniency when evaluating the possible merit of ideas.

One of course prefers ideas that are more likely to be correct because it is correct answers we are seeking and because economy of research effort is extremely important (there are only so many scientific hours and dollars to be spent on any particular

problem). However, it is often very difficult to make such judgments at the outset and, again, our ultimate knowledge and procedures will be no better than the ideas we have thought of and tested. In the context of discovery, one might say that the only requirement is that the idea or method or sign might work, not that it will or does work, and at least for now the scientist has few or no formal methods for deriving probabilities (although Faust and Meehl (1992) have worked on these and related metascience problems; see also Faust (2006, 2008)).

If anecdotal evidence, case studies, and naturalistic studies of “caught” malingerers are viewed mainly within the context of discovery and not verification, we will be in a better position to benefit from their value in uncovering variables or indicators that may prove discriminatory, or in providing the needed grist for the verification mill. However, when the value of evidence is mainly limited to the domain of discovery, it is helpful to recognize and acknowledge these limitations, just as it is unfair to criticize a researcher whose intent is discovery for failing to meet stringent tests of verification. Often these restrictions and cautions are not limited to anecdotal evidence and its close cousins and are mainly a matter of degree, because research on malingering using more advanced designs also suffers from varying levels of concern about representativeness or generalization. More broadly, to the extent evidence or research designs may generate information of potential value but do not permit informed determinations of generalization, they might be thought of more as an exercise in the context of discovery versus verification.

3 What Is the Nature of the Phenomenon We Are Trying to Measure?

3.1 Fundamental Components

It is not an academic exercise to ask, “What is the true nature of the thing we are addressing when we refer to malingering?” This is not a question of definition, which is not too difficult (and, by itself, often resolves no important theoretical issue). Instead it is a question of proper conceptualization of external (real-world) correlates, and in particular whether we are referring to an artificial conglomeration of attributes and behaviors as opposed to something with taxonicity or internal coherence. How are we to think about the clinician’s task if we do not have a reasonably clear idea about just what it is we are trying to identify? For example, the inferences and conclusions we should draw from data can differ greatly depending on whether malingering or falsification represents a continuum, or if falsification in one domain bears a high versus negligible correlation with falsification in other domains. If plaintiff Jones falsifies an early history of alcohol abuse, how much does this tell us about the likelihood he is also misrepresenting a fall down the stairs? If falsification is minimally related across domains, it tells us little; but if it

is highly interrelated, then knowing that Jones underestimates his drinking by 50% could practically tell us that he fell down three steps, not the six he reported.¹

In conceptualizing what malingering might be, at least two components seem to be required. One dimension involves misrepresentation of one's own health status (defined broadly) and the other intentionality. Whether the clinician wants to become involved in examining both dimensions, and whether the practitioner thinks that intention can be evaluated, are separate considerations from whether intentionality is needed in a conceptualization of malingering, which it almost surely is. For example, we would not want to identify a severely depressed patient who misperceives her functioning in an overly negative way or a patient with a parietal tumor who claims his right hand is not his own as malingerers.

One might also wish to parse intentionality into the subcomponents of purposeful or knowing action and the aim or end that is sought. Pretending to be disordered to obtain an undeserved damages award would not seem to equate with pretending to be sleeping so that one's 6-year-old child does not find out it was her parent and not the tooth fairy who left the dollar under the pillow. Or to illustrate the point with perhaps a more compelling or pertinent example, there is a difference between someone fabricating a disorder in an effort to avoid responsibility for a vicious crime and a crime victim feigning death to save his life. One of the difficulties here is unpacking the ontologic and moral issues. On the one hand, there might well be differences between individuals who fake illness for altruistic or at least neutral reasons as opposed to those who do so for self-gain and despite knowing their actions may harm an innocent individual. On the other hand, such distinctions between honorable and dishonorable reasons for malingering may lack objective grounding and can become rather arbitrary or almost purely subjective. For example, the same hockey player who fakes injury to draw a major penalty may be a villain in the visiting arena and a hero in the home arena, and it does not make much sense to say the justifications for the player's actions change during the flight from Montreal to Toronto. One might contrast this circumstance to a situation in which an individual plans and carries out a brutal murder for monetary gain, is caught, and then feigns insanity.

Some social scientists think that these types of value judgments are arbitrary or irrelevant, but assuredly the courts do not share their views. The normative purpose, or at least regulative ideal, of the legal system is to resolve disputes fairly, and this indeed often involves moral judgments and questions of culpability. Individuals' intended goals or reasons for doing something and the legal/moral correctness of their acts frequently decide the outcome of cases. An abused woman who feigns unconsciousness to avoid physical injury is likely to be judged quite differently than an abusing husband who fakes incapacitation so as to lure his spouse into a trap and harm her, even though both are intentionally faking disorder.

¹To reduce the use of the cumbersome "he or she" or "his or her," we will alternate back and forth or vary references when we refer to gender, including the use of "they."

These value issues involve such considerations as whether there would seem to be a morally just versus immoral reason to malingering; whether the malingerer's motives are altruistic, neutral, or self-interested; and whether the act of deception comes at cost to others or victimizes them. Hence, in considering the dimensions of malingering, one might need to ask not only whether the act of providing false information is intended, but also what the individual intends to accomplish and is willing to do given an awareness of the possible consequences for others. Such judgments may reflect societal perceptions for the most part and in some instances are arguably relativistic. Nevertheless, there may well be an intrinsic, qualitatively different dimension one taps beyond falsification and intention when one looks for differences between individuals who will and will not violate major societal norms or engage in deceit for moral versus immoral reasons. Whatever the case, we will mainly limit our focus here to the first two dimensions of intent and misrepresentation.

In legal cases, there is another element that must be considered, although it does not belong on a list of candidate dimensions for malingering. In tort law, a determination of culpability, and the assignment of damages, often depend not only on the presence and extent of harm but also on cause. Smith may be terribly damaged, but if it is not the car accident but the 20-year addictive history that accounts for lowered scores on neuropsychological testing, then the driver who carelessly hit him may owe nothing for neurocognitive maladies.

A plaintiff claiming brain damage may not need to fake or exaggerate disorder at all to mislead the clinician into adopting a conclusion favorable to her case. For example, the plaintiff can simply try to mislead the clinician about cause by hiding or covering up alternative factors that explain her difficulties. Plaintiffs may also overstate prior capabilities to create a false impression about loss of functioning. Whether these alternative forms of deceit represent a separate qualitative dimension or just another phenotypic variation of a genotype is difficult to say, but there is no question that clinicians desire methods for identifying these sorts of deception as well. In fact, attempts to lead clinicians down the wrong causal path may be one of the most common forms of falsification in legal settings and deserves researchers' careful attention.

A definition of malingering that requires intention does not speak to the position or belief that malingering is or can be unconscious. From a legal standpoint, it is not clear how much of a difference there is between fooling oneself and attempting to fool others. Whether a person should be compensated for a supposed act of self-deception is a matter for the courts and juries to decide, and whether mental health professionals should enter into this particular fray is not easily answered and arguably a matter of not only theoretical viewpoint but also pragmatic feasibility (i.e., is the distinction possible to make, especially at an adequate level of scientific certainty?).

Here, what is being sought or accomplished and its justification may be central, such as whether it is the attention of others, reduction in responsibility, or absence from a stressful job; and if changes in circumstances are connected to the event in question and merit financial compensation. For example, if one somehow is using

an accident as a means for assuming the sick role to solicit care and attention from a generally neglectful spouse and to avoid tedious household responsibilities, it is questionable whether someone else should shoulder the cost. In contrast, suppose a person who must drive some distance to work is struck head on by a drunk driver and suffers a severe and prolonged psychological disorder. The injured party stops driving and becomes more dependent on others for emotional support, including a spouse who views emotional maladies as intolerable weaknesses or laughable excuses for skirting personal responsibilities. The injured individual, who is perfectionist and rigid by nature, also has great difficulty accepting personal or psychological faults. In contrast, physical explanations may be far more acceptable to her and her spouse, and she voices physical complaints and perhaps develops beliefs about physical disorders the accident has caused that help accommodate shortcomings and limitations in her functioning that are causally related to the accident. To highlight the differences in these situations another way, one can ask the old Ronald Reagan question: "Are you better off today than you were yesterday?" It is hard to conceptualize an outcome that allows one to avoid what one wants to avoid and pursue what one wants to pursue and be compensated for it (i.e., in which the array of secondary gains far outweigh losses) as comparable to a circumstance in which more enjoyable or favored activities are discontinued and the less pleasant but essential ones now absorb almost all of the individual's energies.

3.2 Malingering Is a Hypothetical Construct

Malingering is a hypothetical construct. It is not a physical entity or an event in the way we normally think of such things (although it of course has an ultimate physical substrate), both of which are classes of variables that potentially can be reduced to a set of observations. The recognition of malingering (or its various forms) as a hypothetical construct carries with it certain methodological implications. First, it is not directly observable but rather must be inferred from a set of observations. To move from observations to constructs requires what philosophers of science refer to as surplus meaning (e.g., assumptions, theoretical postulates, and methods for relating or interconnecting these components). There is understandable concern about not getting too far removed from the observational base or speculating without constraint whatever the scientific data. However, the notion that to go beyond what is directly observable and infuse meaning is a methodological crime (as, say, Skinner seemed to think) is to disregard the commonplace in science. Scientific fields make broad use of hypothetical constructs (some of which later are discovered to be physically identifiable entities), and there is no direct way to go from a set of observations to theoretical constructs, a fatally flawed notion in the early positivist movement and subsequently acknowledged as a mistake. As is sometimes said, one spends the first half of a basic logic class studying deduction and the second half violating it when studying induction, but in science moving from fact to postulate and theory requires the latter.

The nature of the entities we are studying should shape our methodology. For one, if we are dealing with hypothetical constructs, operational definitions are vacuous. The obsession of some psychologists with this defunct and untenable notion of operational definitions—the remnant of a bad idea, almost universally rejected from the outset in the field in which it was proposed—is puzzling. Do we believe we could properly define such things as “quality of life” or “the best interests of the child” operationally? Do we believe if we develop five ways of measuring temperature that we are measuring five different things? Do we believe if a test contains one question, “Are you introverted?” that introversion is what the Introversion Test measures? What conceptual or scientific issue is resolved if we proceed in such a manner? Essentially none. It is worthwhile to seek clarity of language or definition, but this is different from believing that some important conceptual matter is or can be addressed by developing an operational definition. Unfortunately, a close cousin to overvaluation of operational definitions is proposing diagnostic criteria for identifying malingering that are premature given deficiencies in the scientific knowledge base, particularly when they are applied in legal settings (despite what may be clear warnings and cautions by the creators). (For further discussion of diagnostic criteria for malingering, see the final section on caveats.)

The nature of the entities we are studying and the resultant impact on appropriate methodology for *developing* assessment methods needs to be unpacked from the methods that will be most effective in *interpreting* the results these assessment tools generate. It is easy to conflate the two issues. Even if surplus meaning, inference, and theoretical considerations are essential in the development of assessment methods, this does not mean they will also be essential or important when interpreting the outcome these methods generate. For example, theoretical developments and scientific advances might result in an index that provides a simple cutoff point or probability statement. It is not coincidental or contradictory that Meehl, who together with Cronbach (Cronbach & Meehl, 1955; see Faust, 2004) radically impacted the development of assessment methods by emphasizing construct validity (versus blind or pure empiricism), also did more than anyone else to lay out the advantages of statistical or actuarial decision methods (Meehl, 1954/1996; see also Waller, Yonce, Grove, Faust, & Lenzenweger, 2006). One may maximize effectiveness by emphasizing conceptualization and theory in the development of methods, but relying on statistically based methods to interpret results or predict outcomes. Such interpretive or predictive methods need not be processed through the lens of a theory or mediated by theoretical assumptions about mind or behavior. It is commonly just assumed that if methods rest on theory or conceptualization that interpretation of the resultant output should also be based on theory or understanding, but there is no logical reason to form this link. We may need advanced theories of biochemistry to develop markers of certain diseases, but the result may be a test that yields an output that can be interpreted using a simple cutoff score. There is a related common but unwarranted assumption that the nature of the thing being appraised and the form or characteristics of measurement should resemble one another, a matter to be taken up momentarily.

3.3 Distinguishing Between the Nature of Entities and Effective Measurement Strategies

Anyone with at least a dash of scientific realism would likely agree that measurement should ultimately be dictated by external reality; that is, measurement is intended not to construct but rather to reflect what is out there. Therefore, what malingering is and is not will have major impact on the success of different approaches for measuring it. To illustrate the interrelationship between ontology (the nature of things) and measurement, if malingering truly represents multiple dimensions that are largely independent of one another as opposed to a few core characteristics with strong associations, the features of effective assessment tools will likely differ.

It would seem that we encounter an obvious circularity at this point. Measuring devices should fit the nature of malingering, but we do not yet know the nature of malingering and need effective measurement to obtain this knowledge. Hence, it would appear that we need to know more than we know if we are to learn what we need to learn. Under such conditions, how can we proceed? Here again, pseudo-positivism or operationalism will only confound the problem and not get us very far.

Within science (and within the course of human development for that matter) we often encounter this dilemma of needing to know more than we know in order to progress, and yet we frequently find some way around it. In science, this often involves some fairly crude groping around in the dark and a good deal of trial and error (Faust, 1984). We can usually determine whether we are getting somewhere by examining classic criteria for scientific ideas, such as the power to predict and, most importantly and globally, the orderliness of the data revealed (Faust & Meehl, 1992; Meehl, 1991). A phrase like “orderliness of the data” might seem vague and circular, but it has clear conceptual implications among philosophers of science and is probably the most generally accepted criterion for evaluating theories. Circularity, although indeed present, is not that problematical so long as it is partial and not complete (see Meehl, 1991, 1992). The relation between knowing the nature of malingering and measurement is dialectical—the development, ongoing evaluation, and modification of malingering detection devices ought to be based on what we come to know about malingering (our ontological knowledge), whereas our capacity to learn about malingering depends on the state of our measurement tools (our methodological or epistemological competence). Hence, knowing or attempting to know what malingering is and measuring or attempting to measure it necessarily proceed in mutual interdependence.

Although the nature of entities impacts powerfully on the success of different measurement approaches, there is hardly a one-to-one relationship between them. There is often a tendency to conflate ontological and epistemological issues. Ontological claims involve beliefs about the nature of the world or what exists, and epistemological claims involve beliefs about methods for knowing or for learning about the nature of the world. To what extent ontological claims dictate epistemological positions in an idealized system or whether the two should parallel each

other is not a simple matter. However, in the practical world the two need not be isomorphic and can differ or diverge considerably without creating problems, despite what intuition or common sense might seem to suggest. For example, although the entities we intend to measure may be highly complex, this does not necessarily mean useful measurement of them must take complex forms. A few or even a single distinguishing feature may serve to identify a complex entity or condition with considerable accuracy, and at least in the short-term there may be little basis for using complex or multidimensional measurement, especially if the latter is premature and thus relatively ineffective.

Similarly, gross simplification may come very close to reflecting nature accurately (e.g., conceptualizing planetary motion as an ellipse). One might think that because the human brain and mind are complex, prediction must necessarily take into account that complexity and a myriad of data. It may be true that maximizing predictive accuracy ultimately requires that many or all of these complexities are captured, but at present the attempt to do so may create more noise than true variance and make things worse than more simplified approaches. For example, either using past behavior to predict future behavior, or merely predicting that someone will do what most people do, may work far better at times than detailed psychological assessment that attempts to appraise many characteristics or provide deep insights into a person's psyche. Assumptions about features of the human psyche (e.g., that it is complex and involves multidimensional interfaces)—or, more on point, about malingering—do not necessarily dictate measurement that mirrors these features in order to achieve the highest level of accuracy under current conditions.

Given the state of our knowledge at present and perhaps for years to come, there are times that simplifying approaches work as well or better than more complex attempts at measurement, because the latter have limitations that may introduce more error than true variance or dilute stronger predictors by including weaker ones (see the later section on attempting to integrate all of the data and the noncumulative nature of validity). Additionally, deeper understanding of phenomena or causal mechanisms may lead to the development of more sophisticated measurement approaches with decreased or minimal surface resemblance to the things being measured. Who ever imagined that the color of fluid in a tube could tell us whether someone is pregnant, that enzymes might reflect cardiac compromise, or that faint radio signals might provide critical information about the origins of the universe? Thus, the prospect that statistical frequencies might facilitate conclusions about malingering, sometimes much more so than other forms of measurement or understanding, should not lead to premature or reflexive rejection, nor to consternation. Given the importance of what we are trying to accomplish, we should embrace advances whether or not they fit our preconceptions or cognitive aesthetics.

A related questionable or fallacious belief about isomorphism, which was briefly addressed above, is that prediction must be generated by theory or understanding. One can believe that construct validity and conceptual understanding are often indispensable in test development, yet also maintain that highly effective use or application of measures can be largely atheoretical. There is a *massive* literature on

prediction in psychology and related fields showing that statistically based decision procedures almost always equal or exceed clinical judgment and thus are superior overall (see Dawes, Faust, & Meehl, 1989; Faust, Ahern, & Bridges, 2011). If theory or understanding is so essential in reaching conclusions or generating predictions in psychology, then many of these studies should have come out otherwise, especially considering that, once developed, the application of statistical prediction is formulaic and not theory driven or derived. (This is distinct from arguing that good judgment in the selection, use, and application of such methods is not needed, which it is.)

Psychologists who do not distinguish between approaches for developing and appraising tests versus methods for applying them or generating conclusions will often raise ideological arguments that fail to intersect with pragmatic outcomes. For example, in many circumstances heterogeneous measures are better predictors than narrow or more homogeneous measures. A neuropsychological measure that requires multiple functions simultaneously will tend to be much more sensitive to brain damage than one that taps narrower or select capacities, although one may learn little about the specific areas of difficulty involved. If the immediate clinical task is to determine whether brain damage (or dementia, malingering, or some other particular condition or outcome) is present or likely, the selection of the heterogeneous scale might be far and away the most effective and hence the best choice. However, if one adheres doggedly to the notion that prediction should start with understanding or theory, a scale with a diverse mix of items might seem like something to be avoided assiduously. Another but converse form of ontologic-epistemologic isomorphism is to take an atheoretical approach not only to prediction but also to test development and appraisal (as hard-core behaviorists or empiricists once commonly did), something that some strong medicine from Cronbach and Meehl (1955) went a long way toward alleviating. In summary, unwarranted assumptions about ontological and epistemological isomorphism can unnecessarily restrict and impede our efforts to improve measurement.

As follows, the nature of malingering and its relation to needed or preferable measurement approaches may deviate from common belief or expectation. For example, if malingering is a category, one might falsely assume it cannot be identified by scales measuring the amount or extent of some quality (i.e., quantitative standing). However, imagine we were trying to determine whether animals fit the category of zebra. Suppose someone developed a formula that calculated the proportion of white (W) to black (B) and the proportion of white plus black to color of any type (C). If $W:B$ and $W + B:C$ both fall within certain ranges, the animal is to be classified as a zebra. In fact, depending on the animals being considered, such a quantitative index might work rather well, perhaps exceeding 90% accuracy. In turn, despite being based on these relatively isolated, phenotypic characteristics, the ability to identify or classify zebras with a high level of accuracy might then provide a foundation for productive research on the animal and the development of a considerable knowledge base. With a new animal, if one merely calculated the formula, the result might indicate that this knowledge base likely applied (because one was dealing with zebra), in turn permitting one to tap into a good deal of useful

information or predictive power. It might take years for scientists to come up with a clearly superior method of identification, but meanwhile this quantitative procedure, an exercise in approximation or oversimplification, could serve a very useful purpose. We might finally note that effective classification rules, or even knowing whether they are effective, often follows the reverse order, that is, they come after the development of fairly extensive knowledge rather than precede it.

Key Questions About the Nature of Malingering

At present, the key ontological question seems to be whether, at the one extreme, the phenotypic variations of malingering reflect a few basic, interrelated dimensions that have substantial consistency across situations, persons, and falsified conditions or whether, at the other extreme, we are dealing with multiple independent dimensions and loose conglomerations of behaviors that change depending on the person, situation, and condition being feigned. (If we had to place our bet, it would be that malingering consists of multiple distinct categories that may or may not co-occur, and that in addition there are also dimensions of exaggeration or falsification that are not categorical.) Moving from ontology to epistemology, a key measurement issue is the development of methods that, to the extent possible, retain discriminatory power across persons, situations, and variations of falsification, and under conditions in which examinees learn their underlying design. Finally, we consider the key interface between conceptual and measurement issues to be the clinical discriminations of greatest relevance, which are those that the practitioner is required to make but cannot easily accomplish.

If malingering does have at least two basic components, falsification and intentionality, with more than minimal independence from one another, it follows that we need to capture both to identify malingering properly. Furthermore, as we will take up in detail later, any satisfactory method for identifying malingering must account for not only the presence and degree of malingering but also the presence and degree of true injury. To state the obvious, malingering and true injury are not mutually exclusive but can co-exist and are partly independent of one another. Sometimes it is one *versus* the other, but other times it is one *and* the other. If we lose sight of the fundamental difference between opposing and conjoint presentations, research in the area will never approach its true potential and will fail to address pressing legal, social, and moral needs. We contend that one of the largest and most important gaps in our scientific knowledge about malingering involves such combined presentations.

In the original version of this work (Faust & Ackley, 1998), we emphasized the value of taxometric analysis (Meehl, 1995, 1999, 2001, 2004, 2006 [specifically Part IV]; Waller & Meehl, 1998). These methods, which require modest to relatively large samples, serve to clarify the latent structure of variables and are well suited for work on malingering. In addition, even absent definitive or near-definitive methods for identifying group membership (e.g., those malingering versus those not malingering), the methods provide means for identifying optimal cutting scores and

estimating base rates. There has been a gradual increase in the use of taxometric methods in malingering research, and it has sometimes supported the existence of distinct categories (as opposed to underlying dimensions) (e.g., Strong, Glassmire, Frederick, & Greene, 2006; Strong, Greene, & Schinka, 2000) and sometimes has not (e.g., Walters et al., 2008; Walters, Berry, Rogers, Payne, & Granacher, 2009). We think expanded work with such methods promises to add much to our knowledge about categorical versus dimensional status and classification.

Finally, attempts to examine the categorical status of malingering should avoid artificial constraints on its manifestations. Many malingering studies present subjects with only a few measures or options. Although there is nothing wrong with this per se or when conducting certain types of studies, restrictive response options can create fatal problems when one is trying to capture the nature or structure of malingering. In the clinical situation, a potential malingerer has a wide range of options and is almost never forced to fake on a predetermined, narrow range of tests. Rather, the malingerer can fabricate history and symptoms and may well be selective in faking test performances. If the researcher severely restrains the range of options for malingering and forces the individual to fake on a specific or narrow set of measures, a very distorted picture of malingering may emerge. It would be analogous to attempting to determine the underlying characteristics of the dolphin's sensory system by solely measuring whether sound can be detected at a certain level, or to examining the works of Robert Frost by only counting the average number of words in a sentence. None of this should be confused with an argument for considering or integrating all possible evidence in assessing malingering (which is often counterproductive advice; see Faust, 1989, and subsequent material in this volume). Rather, an attempt to determine underlying structure should provide the opportunity for the phenomenon to manifest itself as it is and should not artificially, and severely, constrain its expression.

4 Clinical Needs and Research Agenda

Recognition of the noteworthy gains made in malingering detection should not obscure the considerable challenges that remain. Rather than accept our current tools as good enough and think that, even if there are gaps in research, clinical experience and judgment can almost invariably overcome remaining limitations, we can ask what the most pressing research needs might be. It seems sensible to argue that, all else being equal, the cases that remain most difficult to detect or classify set the main clinical agenda, which in turn sets the main research agenda. Although essentially going hand in hand, such research should also focus on improving or augmenting the best measures and methods, or extending their reach, and not on creating more methods or approaches with validity, but that offer no particular advantages over currently available methods. As straightforward as this seems, a large volume of research may not be directed precisely toward the most pressing clinical needs.

Given the scientific advances that have occurred, a certain percentage of cases are now easily identifiable and can be classified with considerable accuracy. However, in many other instances the clinician's task remains challenging, and more advanced research knowledge and appraisal methods are needed. These remaining difficulties may be obscured or underappreciated exactly because much research does *not* examine these more challenging (but common) presentations and thereby can yield a misleading picture of overall efficacy. Whittling down the percentage of remaining ambiguous or difficult cases will almost surely become progressively more trying and will likely require protracted effort. As we gain more success, those individuals who remain difficult to identify are generally harder and harder cases, and thus the scientific challenges increase accordingly.

There is obviously minimal need for additional research on the types of cases we can identify almost flawlessly. We seemingly should concentrate instead on those cases that frequently exceed our current capacities or knowledge. In general terms, the latter sorts of cases are often those for which there are reasonable grounds to suspect malingering, and one must make the distinction between those who are suspected of malingering and are malingering versus those suspected of malingering who are not malingering. This differentiation is usually far more difficult than distinguishing between cases in which there is almost no reason to suspect malingering versus those in which the evidence for malingering is overwhelming. Yet research is often conducted with these easily identified groups. How informative is it to study very distinctive groups we know how to identify with near certainty in order to learn how to identify those we do not know how to identify (precisely because they lack the distinguishing features of the easily identified groups)? Viewing the main research agenda as cutting into the percentage of difficult to identify or ambiguous cases, we will first discuss the groups of greatest interest, then cover factors that may contribute to false-negative and false-positive errors, next compare clinical needs to common research strategies, and finally present a series of research suggestions.

4.1 Framing the Problem

There is almost nothing more important for advancing malingering research than to identify representative samples of cases. Were this possible, it would greatly facilitate efforts to uncover distinguishing features, such as the characteristics that separate individuals for whom there is a good basis to suspect malingering and who are and are not malingering, and go a long way toward deriving accurate base rate information. In pursuing such aims, it helps to clarify the groups of interest or the individuals who make up the relevant population or subgroups. Figure 2 reflects an attempt to frame this population.

The focus of Fig. 2 is on litigants. We realize that falsification or malingering is not of concern solely in legal cases, but given the main aim of the current text and volume, Fig. 2 is directed toward forensic groups. Further, the characteristics of

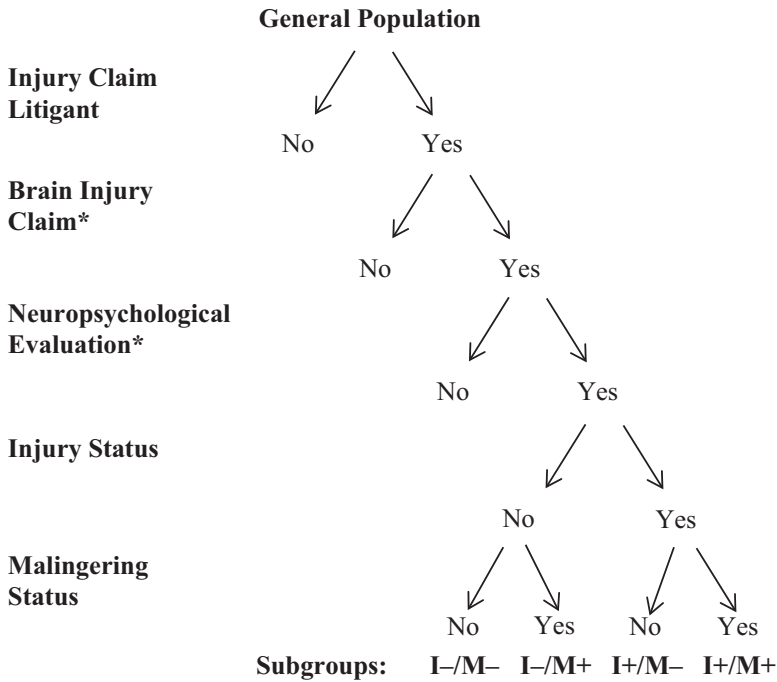


Fig. 2 Identifying relevant subpopulations in forensic neuropsychological evaluations. (“The order of these steps can be reversed and outcome of the neuropsychological evaluation can influence whether a brain injury is claimed.”)

litigants are of greatest relevance for neuropsychological evaluations and research within that arena. The materials that follow are arguably narrower than the overall legal domain because most of our commentary is related to the civil arena, and there are probably important differences or distinctions between civil and criminal cases. For example, in a civil case a major issue may be the occurrence of brain injury and its future consequences. In a criminal case the main concern may be retrospective, such as whether months earlier during a murderous act the defendant’s pre-existing brain injury impeded the capacity to form criminal intent or control behavior at the time of the crime.

As illustrated in Fig. 2, some litigants will claim brain injury or dysfunction, and some of this group will be seen for neuropsychological evaluation. In some cases brain injury is suspected but does not become an element of the case until a neuropsychological evaluation generates an abnormal result. The evaluation may have been initiated in the context of treatment or arranged by an attorney. For the moment, the main point is that, for the neuropsychologist, the overall group of interest is not litigants as a whole or all litigants claiming brain injury but litigants who may or will claim brain injury and who are being evaluated by a neuropsychologist. The importance of all this is that information about the other groups, such as all litigants, will usually be of little or no relevance to research on malingering detection within

neuropsychology in the legal or civil context. Whether the base rate for malingering is, say, 5% or 25% for litigants overall, it matters not a whit because that is not the group that neuropsychologists evaluate in the legal context, and it is the base rate of malingering in the latter group that matters. Similarly, when one thinks about a representative sample or the subcategory from which to try to derive such a sample, the relevant group is not litigants overall, but it is those litigants that neuropsychologists evaluate.

As critical as it is to distinguish between the subgroups in Fig. 2 and their relevance to clinical and research efforts, this figure is an exercise in oversimplification. For example, for the entry, *Neuropsychological Evaluation*, there may well be differences among individuals examined by a treating neuropsychologist, the plaintiff's neuropsychologist, the defense's neuropsychologist, or across two or all of these contexts. Possible distrust of the "opposing" neuropsychologist could lead to systematic differences in evaluation results on average. Furthermore, no attempt has been made to distinguish between such factors as the magnitude or type of injury, the potential presence of co-occurring or independent conditions, the amount of money at stake (e.g., \$50,000 versus \$10,000,000), or litigants' sociodemographic characteristics. There may also be regional differences and differences based on the type of claim or forum (civil, criminal, family court, adult versus juvenile). The mixture of individuals can also change over time. For example, the frequency of cases in which mild brain injury is being claimed can change over the years for a number of reasons (e.g., perhaps a few lawyers have highly visible success with such cases, certain kinds of cases repeatedly bring poor results, or awareness of mild head injury increases due to media and medical attention to war-related or sports-related concussions).

Given these complexities, when a specific base rate is cited for malingering one wonders about its basis, merits, and value, in particular because *general* base rates are often of little help and, rather, one seeks base rates that are narrower and more specifically applicable. To illustrate the point, the base rate for Alzheimer's disease for the overall population is much less helpful than the base rates for a group whose age is comparable to that of the patient, especially if one is dealing with a 7-year-old versus a 70-year-old patient. (The importance of using base rates that are as narrow as possible is discussed later.) The more one considers these sorts of complexities and their implications, the more apparent it becomes that we have often just brushed the surface of clinical and scientific issues crucial to this area.

The flow chart depicted in Fig. 2 is obviously limited to coarse groupings, although in many circumstances even such broad separations may be missed, potentially dooming attempts to get at greater specifics almost before one gets started. The rows labeled *Injury Status* and *Malingering Status* do not reflect a temporal or diagnostic sequence or hierarchy. Rather, they are separated in the flow chart to distinguish them conceptually. We wish to avoid what sometimes seems to be a "versus" bias in this area, or the tendency to treat these categories as if they were exclusive of one another more often or to a greater degree than is warranted. Combining injury status and malingering status, we end up with four subgroups (i.e., not injured and not malingering; not injured and malingering; injured and not