Clínica y Salud 24 (2013) 129-138

Clínica y Salud

www.elsevier.es/clysa

Symptom validity assessment in European countries: Development and state of the art

Thomas Merten[a]*, Brechje Dandachi-FitzGerald[b], Vicki Hall[c], Ben A. Schmand[d], Pablo Santamaría[e], and Héctor González-Ordi[f]

[a]*Vivantes Netzwerk für Gesundheit, Klinikum im Friedrichshain, Berlin, Germany*
[b]*Maastricht University, The Netherlands*
[c]*Forensic Psychology Practice, Sutton Coldfield, West Midlands, Great Britain*
[d]*University of Amsterdam, The Netherlands*
[e]*TEA Ediciones, Madrid, Spain*
[f]*Universidad Complutense de Madrid, Spain*

ARTICLE INFORMATION

Manuscript received: 10/07/2013
Revision received: 31/08/2013
Accepted: 15/09/2013

Keywords:
Symptom validity assessment
Malingering
Performance validity tests
History

ABSTRACT

In the past, the practice of symptom validity assessment (SVA) in European countries was considerably lagging behind developments in North America, with the topic of *malingering* being largely taboo for psychological and medical professionals. This was being changed in the course of the past decade with a growing interest in methods for the assessment of negative response bias. European estimates of suboptimal test performance in civil and social forensic contexts point at base rates similar to those obtained in North America. Symptom over-reporting and underperformance in neuropsychological examinations appear to occur in a sizable proportion of patients. Although there is considerable progress in establishing SVA as an integral and indispensable part of psychological and neuropsychological assessment in some countries, others appear to lag behind. In some countries there is still enormous resistance against SVA from part of the neuropsychological and psychiatric communities.

Evaluación de la validez de los síntomas en Europa: evolución y situación actual

Palabras clave:
Evaluación de la validez de los síntomas
Simulación
Tests de validez de ejecución
Historia

RESUMEN

Hasta no hace mucho tiempo la evaluación de la validez de los síntomas en Europa, tanto en su vertiente científico-académica como en la práctica profesional, estaba muy por detrás de los avances que se producían en Norteamérica y particularmente la simulación se consideraba un tema tabú entre los profesionales de la Psicología y la Medicina. En la última década las cosas parecen haber cambiado, observándose un incremento en el interés por la evaluación del sesgo de respuesta negativo. Las tasas base de prevalencia obtenidas utilizando pruebas de rendimiento subóptimo en contextos civiles y forenses son similares a las obtenidas en Norteamérica. Los fenómenos de exageración de síntomas en autoinformes y rendimiento insuficiente en pruebas neuropsicológicas parecen ocurrir en semejante proporción de pacientes. Aunque se han producido avances notables en el establecimiento de la evaluación de la validez de los síntomas como una parte integral e indispensable de la evaluación psicológica y neuropsicológica en algunos países europeos, en otros sin embargo la situación es mucho más incipiente. De hecho, en algunos países sigue existiendo una gran resistencia a la evaluación de la validez de los síntomas proveniente de algunos profesionales de la psiquiatría y la neuropsicología.

*Correspondence concerning this article should be sent to Thomas Merten. Vivantes Netzwerk für Gesundheit. Klinikum im Friedrichshain. Dpt. of Neurology. Landsberger Allee 49. D - 10249 Berlin, Germany. E-mail: thomas.merten@vivantes.de

1130-5274/$ - see front matter

DOI: http://dx.doi.org/10.5093/cl2013a14

Symptom Validity Assessment

In the past 20 years, symptom validity assessment (SVA) has become a leading topic in clinical neuropsychology, with few areas surpassing it in terms of the interest and controversy that it has attracted as well as the proliferation of published research papers. Neuropsychological assessment has always maintained a close link to standardized testing, which makes clinical neuropsychology a data driven discipline and, at the same time, presents the greatest challenge to the validity of individual diagnostic decision making. Neuropsychological test results are highly dependent upon the willingness of the patient to employ optimal test effort, and neuropsychologists are usually not capable of determining the validity of test data on the basis of clinical intuition (e.g., Heaton, Smith, Lehman, & Vogt, 1978).

The era of modern SVA began in the late 1980s, when standardized forced-choice symptom validity tests (SVTs) were introduced in the context of a renewed interest in malingering research. In the following 25 years, neuropsychologists have maintained a leading role in the development of methods and in conceptual refinement in what may be described as the differential diagnosis of non-authentic symptom production and symptom report (e.g., Carone & Bush, 2013; Sweet & Guidotti-Breting, 2013).

In line with Greve, Bianchini, and Brewer (2013), the term SVT is used in the context of this paper as a superordinate term for all methods that tap negative response bias both in symptom production (including underperformance in psychological tests) and symptom over-reporting. For cognitive SVTs determining possible underperformance, Larrabee (2012) has recently proposed the term performance validity tests (PVTs), while those SVTs determining the validity of self reported mental or cognitive impairment may be called self-report validity tests (cf. Figure 1).

The present paper focusses primarily on the current state of PVT usage, the traditional field of SVA, while self-report measures of symptom validity will not be discussed in detail. The tradition of self-report validity scales is far more complex and has had a long tradition in personality research, clinical psychology, and various branches of applied psychology.

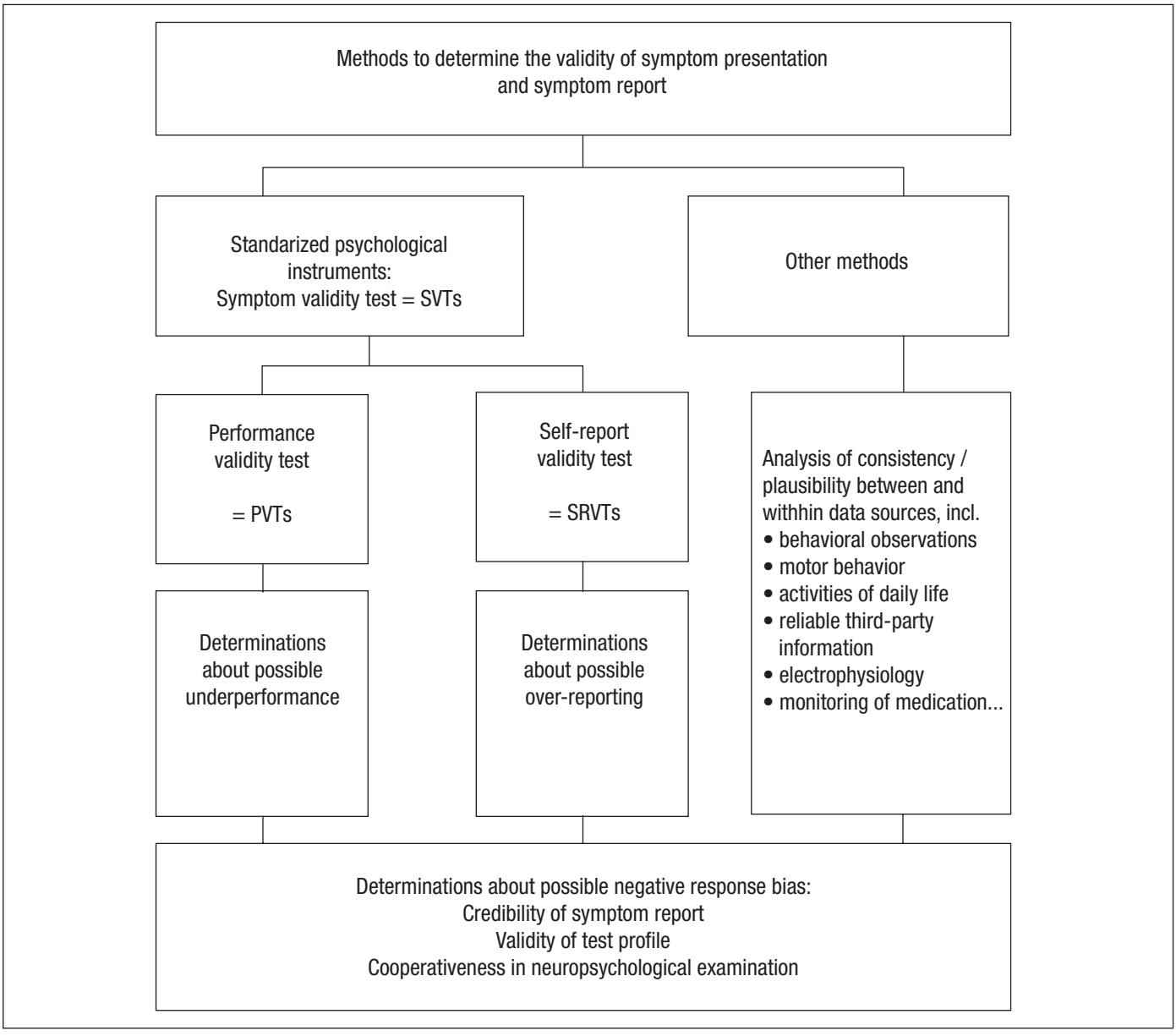

Figure 1. System of methods to determine symptom validity.

Short historical survey of symptom validity assessment in Europe

The problem of malingering, in terms of its definition and detection, has deep historical roots, which shall not be traced here in detail. In the 1811 edition of the Classical Dictionary of Vulgar Tongue, first published in 1785, a keyword "King's Bad Bargain" defines as one of those bad bargains "a malingeror, or soldier, who shirks his duty" (Grose, 1811). For civil society, the 19th century history of the railway spine (Erichsen, 1867; Thomann & Rauschmann, 2003) has fully demonstrated the difficult conceptualization of bizarre symptom presentations in the context of liability claims. Attempts at medical theory building trying to conform to patient behavior may, from today's perspective, appear no less bizarre than the symptomatology they aimed to explain. Yet, unsolved questions as to a valid differential diagnosis of intentional vs. unintentional symptom distortions continue to haunt the medical and psychological literature (cf. Merten & Merckelbach, 2013b).

The Franco-Swiss psychologist André Rey (1941, 1958) pioneered a field which we conceptualize today as SVA. He introduced a number of instruments aiming to detect invalid test performance, such as the Fifteen-Item Test (FIT), the Dot Counting Test (DCT) or the Word Recognition Test (WRT). While these tests perform poorly in contrast to modern PVTs (e.g., Hartman, 2002), Rey (1958) took a rather modern stance in that he warned not to base far reaching diagnostic decisions solely on the results of a single test, but to include multiple information such as the individual constellation of the case, personality, and motivation (cf. Frederick, 2002, for a sound historical appraisal of Rey's contribution to SVA). There were other approaches to tackle the problem of malingering detection which, in modern terminology, would be called *embedded measures* (that means applying empirically derived cut scores or indices based on standard test results to serve as an indicator of symptom validity). For example, before leaving Germany for British Columbia, Otfried Spreen worked on measures to identify feigned cognitive impairment within the Visual Retention Test (Benton, 1945; Benton & Spreen, 1961). Incidentally, Spreen (1963) was also the author of the first German MMPI version. Gudjonsson and Shackleton (1986) later generated another embedded indicator of response validity from the Standard Progressive Matrices.

Forced-choice PVTs were developed in the United States in the 1960s. This methodological approach, which allows for determinations of intentional response manipulations in cases of below-chance performance (Merten & Merckelbach, 2013a, for a recent review), prepared the transition into the modern era of SVA. There was an early European attempt to develop a forced-choice approach for the detection of invalid performance in Benton's (1945) Visual Retention Test. However, this work by Loewer and Ulrich (1971) has largely remained unnoticed, probably because it was published in German language. Thus, it is a paper by Pankratz (1983) that is usually conceived to mark the introduction of forced-choice testing in the field of examining possible malingering in cases of claimed memory impairment. Although he also authored an article in German language (Pankratz & Paar, 1988) there was no detectable repercussion on assessment practice in the German speaking countries.

SVA began to be used and discussed on a broader basis by American neuropsychologists in the early 1990s. During this time they developed a wealth of publications and a range of standardized, commercially available tests, whilst the whole of Europe appeared to be lagging behind. The developmental delay can be estimated to amount to approximately a decade for the most advanced countries (in terms of SVA); in other countries, the delay appears to be much longer still. Yet, a number of PVTs were developed in Europe in the 1990s, such as the Coin-in-the-Hand Test (Kapur, 1994), the Amsterdam Short-Term Memory Test (Schmand, de Sterke, & Lindeboom, 1999), and the Bremen Symptom Validation (Heubrock & Petermann, 2000). A number of PVTs were adapted to European languages, with a multilingual computerized version of the Word Memory Test (WMT, Green, 2003) being the most prominent of them.

Also in the 1990s, some reviews (e.g., Heubrock & Petermann, 1998) and the first empirical studies were published (e.g., Schmand et al., 1998) marking the beginning of a broader research interest in symptom validity assessment. In North America, the term *malingering research* was used in the beginning, and symptom validity assessment primarily centred on the determination of possible *malingering*. In this vein, several instruments have this term in their names, like the German-language Aggravations und Simulationstest, AST (Exaggeration and Malingering Test, Eberl & Wilhelm, 2007). It was only after the turn of the millennium that a conceptual clarification took place on a larger scale. Today, as Merten and Merckelbach (2013b, p. 122) summarized, most experts in SVA would agree that "a) SVTs may help to clarify the nature of certain symptom constellations; b) symptom validity assessment comprises both self-report measures that tap over-endorsement of symptoms and tasks (i.e., 'effort tests') that tap cognitive underperformance (Larrabee, 2012, referred to these cognitive SVTs as performance validity test); c) symptom over-endorsement and/or cognitive underperformance represent two aspects of negative response bias: in some cases they occur together, in other cases only one of the two aspects is present […]; d) malingering is considered to be only one possible source of negative response bias."

It was only with the beginning of the third millennium that European SVT research was performed on a larger scale and symptom validity assessment began to be introduced on a larger scale into practical work of neuropsychologists. First workshops were organized in some countries (as in Britain or The Netherlands), raising immense interest. In London, a post-qualification training day on symptom validity, organized by the British Psychological Society Division of Neuropsychology in February 2006, attracted sizable audience. One of the speakers, Dr. Brooks, "described the meeting as a turning point in British neuropsychological practice" (Alcott, 2006, p. 1).

A landmark was reached with the First European Symposium on Symptom Validity Assessment, held in Würzburg, Germany, in 2009, followed by updates in London, UK, in 2011 (Anderson, 2010; Plohmann, 2011) where major SVT researchers presented. Finally, the third symposium, held in Wuerzburg again, in 2013, fully focussed on European developments and perspectives (Plohmann, in this issue). The fourth symposium is planned for Maastricht, The Netherlands, in two years' time from now.

The bulk of published European research on SVA is authored by researchers from only four countries: Germany, Great Britain, The Netherlands, and Spain. Only few published studies are authored from psychologists of other nationalities, such as Austria (e.g., Schiemann, 2003), Portugal (Martins & Martins, 2010; Simões et al., 2010), or Switzerland (Giger, Merten, Merckelbach, & Oswald, 2010) while there is no or almost no evidence for substantial research activities in some major European nations like France or Italy. This might reflect large differences in the degree of acceptance that symptom validity assessment has reached within Europe.

In the following paragraphs, we will describe the reasons why valid determinations about the genuineness of personal injury and mental health claims are increasingly important and call for reliable methods to distinguish between legitimate and illegitimate claims. Furthermore, the results of a recent survey among neuropsychologists are reported, and summaries of the developments in four European countries will be given.

Claims culture: The British example

In many European countries, public compensation schemes, social security, and social welfare systems have come to the limits of their capacity, and some branches of the private insurance industry are aching under the burden of personal injury claims and the cost

of mental health care, with an ever increasing number of minor mental disorders to result in long-term sick leave and retreat from work. Although the public appears to be well aware of abuse and fraudulent claims and scores of spectacular fraud cases have been published (e.g., Frei, 2004; as an example see the incredible reports from Zakynthos, the Greek "Island of the Blind", Squires, 2013), nobody knows the true financial burden of false symptom report and false imputations. For the United States, Chafetz and Underhill (in press) have recently estimated the cost of malingering in selected Social Security programs. For adult mental disorder claimants, the authors estimated the annual cost at about $20 billion. The acuity of the problem shall be illustrated in some more detail by the situation in the United Kingdom.

In recent years, the UK claims culture has become more widespread, increasing the need for performance validity testing. This is reflected in the "disability paradox" showing an increase in disability claims (rise from just under 2.5 million to 3.2 million) whilst the average health status improves (Department of Work & Pensions, 2013). The 2012 statistics of this UK Department show that there were 3.28 million recipients of Disability Living Allowance (DLA) and £1.2 billion of total benefit expenditure is overpaid due to fraud. However, this is likely to be only the tip of the iceberg, as the validity of claims is not routinely screened using standardized tests. Factors which may contribute to malingering or symptom exaggeration include the "benefits trap", which refers to disincentives for returning to work being built into the UK welfare system (i.e., often people with disabilities cannot earn as much as their benefits if they return to work), although reforms to increase work incentives are only just starting to be made to this system.

The growing claims culture is also evident in the litigation system. There was a five per cent increase in the proportion of accidents involving third-party injury claims in England and Wales last year, making Britain more litigious than America and placing it as the "Whiplash capital or the world" (Ministry of Justice, 2013). Government figures show that between 2006 and 2011 the number of reported road traffic accidents fell by 20%. Over the same period, there was a 60% rise in road traffic accident personal injury claims. According to industry figures, the level of undetected fraud is currently considered to cost the industry around £1 billion a year and they claim that of every hour of every day 15 fraudulent insurance claims are exposed in the UK (Association of British Insurers, 2012; Ministry of Justice, 2013). Press coverage suggests that this is due to a "claims manufacturing industry" that has "gone into overdrive" to entice drivers into making easy claims which has in turn increased the prevalence rate of fraud (Massey, 2013).

Prior to the increase in claims culture, the prevalence rate of performance invalidity (measured by failure on the Test of Memory Malingering) in UK head injury litigants referred for neuropsychological assessment was found to be 33% (Moss, Fokias, Jones, & Quinn, 2003). This is fairly consistent with the estimates of malingering and symptom Exaggeration in North America (Mittenberg, Paton, Canyock, & Condit, 2002). As disability benefits and the "benefits trap" may increase the risk of symptom exaggeration and malingering, there is also a clear need to investigate how common this is currently with non-litigating NHS disability cohorts as well as more up to date prevalence rates for groups perusing litigation.

Despite this growing risk of symptom exaggeration and malingering, in the UK there is a problem with using PVTs to diagnose malingering as the unmasking of malingering is seen as a task for the legal process rather than the medical one (Halligan, Bass, & Oakley, 2004). *Malingering* is therefore seen as a taboo subject and most UK clinicians are skeptical about whether PVTs can differentiate "malingering" from other forms of symptom invalidity (e.g., somatoform disorders and iatrogenic symptoms; McMillan et al., 2009). This is contrary to the North American diagnostic criteria for malingering that infer *intent* from the combined improbability of events (i.e., multiple failures on symptom validity indicators; Larrabee, Greiffenstein, Greve, & Bianchini, 2007).

Surveys on Symptom Validity Assessment in Europe

A first survey on symptom validity practice in Europe was performed in Great Britain by McCarter, Walton, Brooks, and Powell (2009). It has shown that only 16% of clinicians working solely in clinical settings employ PVTs for more than half of their assessments, compared to 73% of those conducting medico-legal assessments, but only 59% of respondents in the survey admitted to always or almost always formally evaluating performance validity with litigants using PVTs. The PVTs most commonly used by UK clinicians were the Test of Memory Malingering (TOMM, Tombaugh, 1996) (50% of all respondents), the Rey 15-Item Test (Rey, 1958) (24% of all respondents), and the WMT (Green, 2003) (24% of all respondents). Only 5% of respondents used embedded symptom validity indicators and only 11% of clinicians used self-report SVTs such as the MMPI (McCarter et al., 2009). The UK therefore seems to be behind the North American recommendations for symptom validity assessment, which stipulates that multiple sources of information should be used to detect response bias (Heilbronner et al., 2009), and this is necessary for the neuropsychological criteria for malingering (e.g., Bianchini, Greve, & Glynn, 2005; Slick, Sherman, & Iverson, 1999).

A second survey was performed in six European countries: Denmark, Finland, Germany, Italy, The Netherlands, and Norway (Dandachi-FitzGerald, Ponds, & Merten, in press). In this survey the median estimate of insufficient effort was 10% in clinical and 15% in forensic assessments. The median estimate of malingering (i.e., insufficient effort specifically due to malingering) was 4% in clinical and 10% in forensic assessments. And while at least the forensic estimates of insufficient effort and malingering were substantially lower than empirical studies show, these numbers still illustrate an increased awareness in neuropsychologists in these European countries as opposed to few decades ago, when the general view prevailed that these phenomena were virtually non-existent. Interestingly, the survey revealed substantially lower prevalence estimates of malingering in own clinical and forensic assessments than in general (Figure 2). Neuropsychologists appear to be hesitant to acknowledge the occurrence of malingering in their own patients. Perhaps neuropsychologists are still too much inclined to believe what a patient says and does during an assessment. This stresses the importance of using objective methods for evaluating symptom validity.

So, what are the tools and tests neuropsychologists use to evaluate the validity of the obtained diagnostic data? The five methods most frequently indicated for the determination of symptom validity were (in rank order): discrepancies between records, self-reporting, and observed behaviour; severity of cognitive impairment inconsistent with the condition; pattern of cognitive impairment inconsistent with the condition; implausible self-reported symptoms in the interview; and implausible changes in test scores across repeated examinations. All objective empirically validated methods (e.g., stand-alone PVTs, embedded indicators, validity scales on objective personality tests) ranked at the bottom. Looking at the analyses per society, there were two exceptions: the Dutch respondents (stand-alone PVTs ranked second) and the Norwegian respondents (stand-alone PVTs ranked fifth).

The most commonly used PVTs were: the ASTM, Rey FIT, TOMM, and the WMT.

As might be expected, only a small minority of respondents (12%) indicated always using a PVT in a clinical assessment. Surprisingly, it was also a minority of the respondents that stated they always included a PVT in a forensic assessment (45%). In this respect, the survey showed that considerable differences existed between the

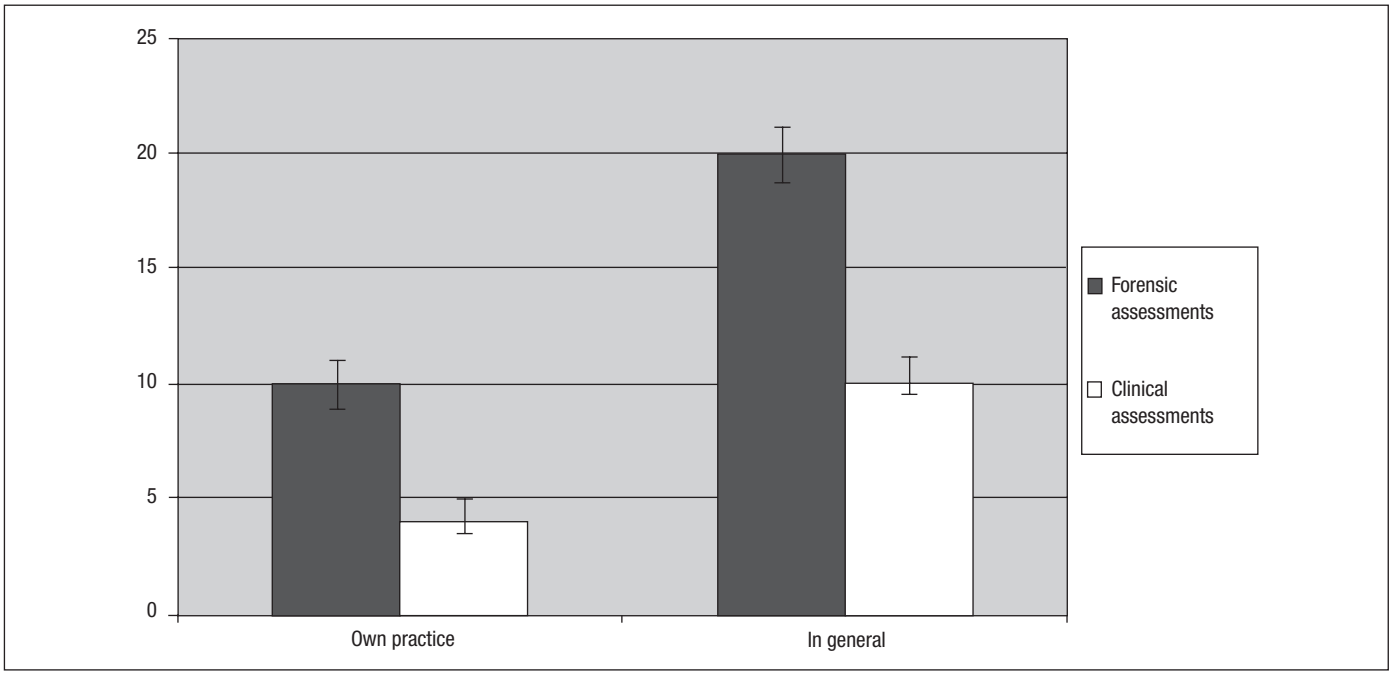

Figure 2. Prevalence estimates of malingering: own versus general estimates in clinical and forensic assessments(median percentages). Interquartile ranges were 19, 5, 30, and 10, respectively.

respondents of the various societies, with lowest numbers in Italy (22%) and Finland (14%) and highest numbers in the Netherlands (70%) and Norway (69%) (Figure 3). This supports the notion that important cultural differences still exist in how neuropsychologists in the various European countries relate to the topic of symptom validity.

In conclusion, on the positive side there is acknowledgement that non-credible symptom reports and malingering occur. The surveys' findings, however, also show that the use of clinical impressions still prevails over the use of empirically validated tests to determine the symptom credibility. Although these subjective methods certainly have their value, it is vital that they are being used in conjunction with empirically validated tests.

Spain

In the last few decades, several research teams and lines have been established in Spain regarding research on malingering and SVA in different domains such as the neuropsychological context (e.g., Vilar-López, Gómez-Río, Caracuel-Romero, Llamas-Elvira, & Pérez-García, 2008), the forensic context (e.g., Arce, Fariña, Carballal, & Novo, 2009; Jiménez-Gómez & Sánchez-Crespo, 2004; Sánchez-Crespo, Jiménez-Gómez, Ampudia-Rueda, & Merino-Barragán, 2012; Zaldivar-Basurto, García-Montes, López-Rios, Molina-Moreno, & Santiago, 2007), the medico-legal context (e.g., Capilla-Ramírez & González-Ordi, 2009; González-Ordi, Capilla-Ramírez, Santamaría, & Casado-Morales, 2012), González-Ordi, Santamaría, & Fernández-Marín, 2010; the military context (e.g., García-Silgo & Robles-Sánchez, 2010), and more generally the study of verbal and non-verbal cues to detect deception (e.g., Masip, Garrido, & Herrero, 2009).

Within these fields of research, empirical results are slowly becoming more and more available for the professionals by means of the publication of journal monograph issues (Lemos-Giráldez, 2005), book chapters devoted to malingering and SVA (González-Ordi &

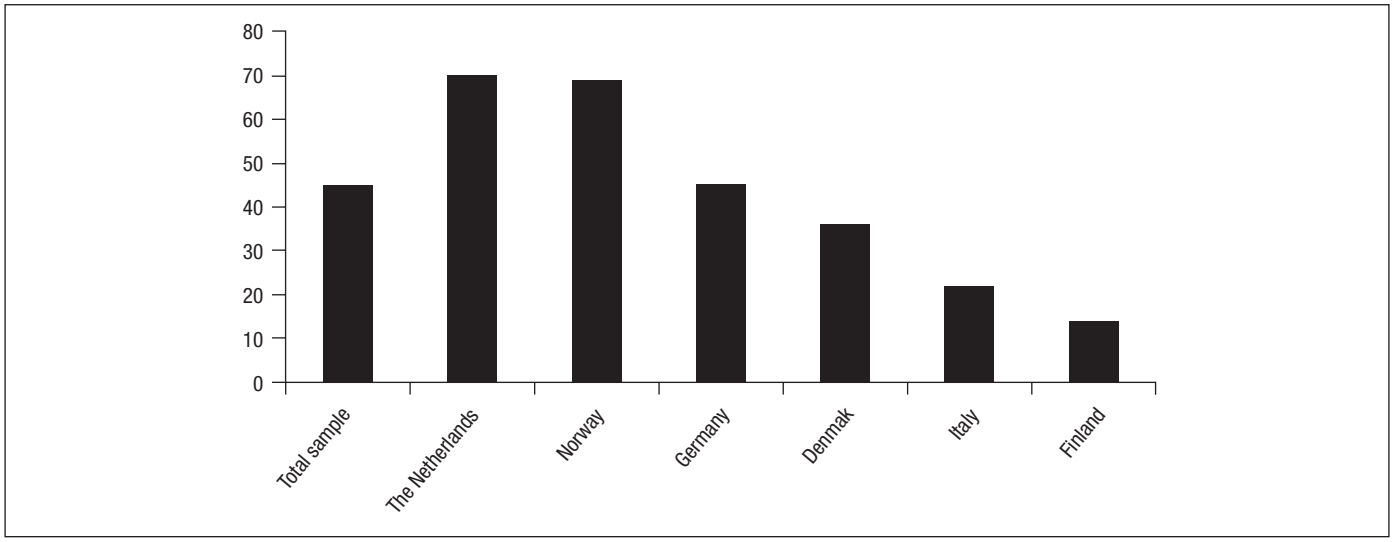

Figure 3. Inclusion of SVT in *every* forensic assessment (% respondents).

Gancedo-Rojí, 1999; Vilar-López & Aliaga, 2009), or theoretical and practical handbooks (e.g., González-Ordi, Santamaría, & Capilla-Ramírez, 2012), as well as continuing postgraduate education through workshops, symposia and conferences, generally in medico-legal and forensic contexts.

Spanish adaptations of the internationally most widespread tests used for the assessment of symptom validity and malingering have fostered scientific, forensic, and medico-legal applications among psychological and neuropsychological practitioners. Remarkably, a significant number of Spanish-adapted symptom validity assessments rapidly became available to clinical and forensic professionals within less than half a decade (e.g., the Spanish adaptation of the Structured Inventory of Malingered Symptomatology – SIMS, González-Ordi & Santamaria, 2009; the Minnesota Multiphasic Personality Inventory-2-Restructured Form – MMPI-2-RFl, Santamaría, 2009; the Personality Assessment Inventory – PAI, Ortiz-Tallo, Santamaría, Cardenal, & Sánchez, 2011; and the Test of Memory Malingering – TOMM, Vilar-López, Pérez-García, & Puente, 2011). Moreover, specific forensic protocols were developed (e.g., Arce & Fariña, 2005).

However, in spite of this promising research panorama there are numerous shadows regarding the professional practice of SVA and malingering assessment. Firstly, there is no reliable data available of applied assessment SVA protocols in different contexts. Furthermore, SVA and the question of possible malingering are not yet regularly taken into account in many forensic and medico-legal contexts. A guide for good practice is still in demand. A general professional debate and discussion have not yet occurred in Spain, as is the case in other European countries, like Germany and Switzerland. Secondly, a risky application of SVTs (like the SIMS) can be identified. This is the case when instruments are used in isolation and without an adequate conceptual and methodological knowledge on the part of the test user who may see a "magic road" to the detection of deception in these tests. In face of this, the role of modern SVA in detection of malingering and the importance of a multidimensional approach in different contexts has to be popularized among professionals (Capilla & Gonzalez-Ordi, 2009; Merten & Merckelbach, 2013b). Multidimensional and multi-method approaches may contribute to avoid inaccurate diagnoses and to decrease the rate of false positive and false negative classification errors, both having a considerable economic and psychosocial impact.

In summary, although there are quality seeds available to get a good harvest in Spain, further research and educational effort will be necessary for establishing sound practice guidelines and protocols, both for researchers and professionals. They should aim to (1) extend a regular use of SVA to any context where negative response biases can occur (clinical, forensic, and medico-legal); (2) adapt SVA to specific disorders (like pain disorder, anxiety and affective disorders, amnesia, posttraumatic stress, sequelae of sexual assault and abuse, etc.); (3) promote a multidimensional approach to SVA and avoid simplified diagnostic decision making based on the isolated use of instruments; and (4) carefully avoid the classification errors (false positives and false negatives).

The Netherlands

In 1997, a Dutch textbook on clinical neuropsychology was published that nicely illustrated the skepticism and controversies on the topic of malingering, which dominated at that time in the Netherlands. This textbook already contained a chapter on malingering, in which the authors advised their colleagues to always administer a symptom validity test when a patient has an incentive to demonstrate cognitive deficits (Schmand & Ponds, 1997). In the same book, the authors of a chapter on the chronic whiplash syndrome stated: "In our opinion, the hypothesis that the desire for financial gain would play a role in whiplash needs not to be taken seriously" (Van Zomeren & Saan, 1997). In those early days, it was believed that malingering was only a problem in the United States and not in Europe, due to the American claim culture with huge financial stakes (e.g., Ponds, de Lugt, Verhey, & Jolles, 1995). Neuropsychologists were reluctant to introduce SVTs, because they felt it as a sign of mistrust towards the patient, opposing their primary task of determining the veracity of the complaints. The use of SVTs implies misinforming the patient, which was considered to conflict with the ethical code of psychologists. As formulated by a prominent colleague: "Psychologists should not allow themselves to be hired to play detective for the insurance companies" (Hofstee, 2005).

From the nineties, SVA has nevertheless been an active research topic in the Netherlands. In 1999, the Amsterdam Short-Term Memory Test, a Dutch PVT was published (Schmand et al., 1999), and English and German versions became available (Schmand & Lindeboom, 2005). The WMT (Green, 2003), MSVT (Green, 2004), and the SIMS (Merckelbach & Smith, 2003) have been made available in Dutch. As in Northern America, the initial focus was on the detection of malingering in forensic assessments with SVTs, in those days often referred to as "malinger tests". However, SVTs were soon applied in clinical settings. Van der Werf, Prins, Jongen, van der Meer, and Bleijenberg (2000) noted that a modest proportion of patients with chronic fatigue syndrome showed non-credible performance. In 2005, a Dutch-Spanish study was published, in which the WMT was used in a clinical sample of schizophrenic patients (Gorissen, Sanz, & Schmand, 2005). Lower scores on the WMT were associated with higher levels of negative symptoms and it was hypothesized that an underlying motivational deficit could explain the WMT failure rate in this sample. These studies are exemplary of the transitional period towards a better understanding of what SVTs actually measure (i.e., the behaviors of cognitive underperformance and symptom over-reporting and not directly the underlying motive), a period marked by great conceptual confusion. The nomenclature expanded and terms such as "suboptimal effort", "cognitive underperformance", and "symptom exaggeration" were being used interchangeably, even as synonyms of the concept of malingering.

The conceptual confusion surrounding SVA has certainly hampered resolving the controversies among Dutch neuropsychologists. The initially assumed dichotomy between true patients and malingerers and the lack of conceptual clarity probably evoked a counter reaction. In recent years, many Dutch neuropsychologists have stressed that cognitive underperformance and symptom exaggeration are not necessarily conscious behaviors nor behaviors motivated by external gains. Instead, emphasis is placed on explanations such as pain, fatigue, and emotional factors (e.g., the ill-defined concept of 'cry for help', struggle for recognition) for symptom invalidity. Currently, the majority of Dutch neuropsychologists acknowledge cognitive underperformance and symptom over-reporting as threats to the validity of the test results and most of them, at least in forensic assessments, use SVTs. However, the idea that symptom invalidity can arise from deceitful behavior of the examinee still seems to be a taboo.

Recent research activities have focused on symptom validity in patients clinically referred for neuropsychological assessments. In a mixed psychiatric sample, 8% of the patients failed two SVTs, the SIMS and ASTM (Dandachi-FitzGerald, Ponds, Peters, & Merckelbach, 2011). The WMT has also been researched in memory clinic patients. Failing the effort indices of the WMT negatively impacted the association between hippocampal damage and memory performance (Rienstra, Groot et al., 2013), thereby obscuring normal brain-behavior relationships. Analysis of the profile of WMT subtests contributed to the prediction of conversion to dementia within 2 years (Rienstra, Twennaar, & Schmand, 2013). Thus, SVTs may even be used to aid the early diagnostics of dementia.

Also, there have been experimental studies into the residual effects of feigning. Volunteers instructed to first malinger

psychological symptoms and then to give up their role and act honestly continued to show heightened symptom levels on re-administration of a symptom list as compared to participants being asked to fill out the list twice in an honest manner. Cognitive dissonance is hypothesized as the underlying mechanism. These studies give a first indication of how malingering may evolve in somatoform presentations (Merckelbach, Dandachi-FitzGerald, van Mulken, Ponds, & Niesten, in press; Merckelbach, Jelicic, & Pieters, 2011).

Great Britain

In Great Britain (GB), the widespread use of PVTs has not quite caught on in clinical practice (McCarter et al., 2009) and many clinicians remain skeptical about their necessity. Despite this lack of popularity, their importance is recognized by the British Psychological Society (BPS), who has written a position paper on performance validity testing, which recommends their use in clinical and medico-legal practice (McMillan et al., 2009).

The majority of British PVT research has focused on test validation but there is a paucity of research, showing that GB's need and interest in PVTs is not reflected in the evidence base. This lack of research is recognized by the BPS, who highlights the need for more British PVT research (MacMillan et al., 2009). The Coin-in-the-Hand Test (CIHT, Kapur, 1994), a bedside PVT, is the only PVT which has been developed in GB and is utilized by clinicians (McCarter et al., 2009). Although the Warrington Recognition Memory Test (RMT, Warrington, 1984) was developed in GB and is used as a test of response bias (Millis, 1994) it was not originally developed as a PVT and there is no British validation data. The CIHT uses chance levels cutoffs and has been shown to have good sensitivity with simulator groups and good specificity with severe cognitive impairment (Hanley, Baker, & Ledson, 1999; Kapur, 1994; Kelly, Baker, Broek, Jackson, & Humphries, 2005; Schroeder, Peck, Buddin, Heinrichs, & Baade, 2012). This worse than chance performance methodology was applied in the very early development of PVTs (Pankratz, 1983), but since this time methodology has become much more sophisticated, given its likely limited sensitivity with "known groups" of clinical malingerers. Nevertheless, the CIHT remains a useful bedside screen.

Although the TOMM and WMT were reported as the most commonly used PVTs in GB, there is only one British validation study for the WMT (Hall, Worthington, & Venables, 2013) and no validation studies for the TOMM. The tests which have been the most extensively validated in GB are the Coin-in-the-Hand Test, the Rey 15 item test (Fisher & Rose, 2005; Kelly et al., 2006), and the Medical Symptom Validity Test (Gill, Green, Flaro, & Pucci, 2007; Singhal, Green, Ashaye, Shankar, & Gill, 2009), although there are still few validation studies with regards to these tests. The continued popularity of the Rey is surprising, as it has been shown to lack sensitivity to malingering and simulation whilst also lacking specificity (Larrabee, 2005; Vallabhajosula & van Gorp, 2001). The remaining British PVT validation studies have tended to be with stand-alone measures which are rarely used (e.g., Autobiographical Memory Test, Jenkins, Kapur, & Kopelman, 2009) or embedded PVTs which again have little popularity (e.g., Multiple Errands Test, Castiel, Alderman, Jenkins, Knight, & Burgess, 2012; Wechsler Memory Scale-Third Edition, Mental Control Test, Kelly et al, 2006; Wechsler Memory Scale, Word List Recognition, Hacker & Jones, 2008). The British studies have generally been in relation to non-litigant neurological groups, such as traumatic brain injury (Wogar, van den Broek, Bradshaw, & Szabadi, 1998) and dementia (Singhal et al., 2009) and one study applying PVTs to medically unexplained symptoms (Kemp et al., 2008). The lack of litigant studies (with the exception of Moss et al., 2003) is surprising given that this is the area which is most applicable to PVTs, given the high prevalence rate of symptom exaggeration in litigant groups. There is also a lack of British PVT studies with chronic pain or in the context of criminal forensic neuropsychological assessments and this is also reflected in clinical practice with few neuropsychologists providing these specialist assessments. Interestingly, all of the British studies have used simulation validation paradigms and there are no British known-groups studies with clinical malingerers. This trend to use simulation paradigms is behind the North American literature, which has shifted from simulation designs to known-groups studies, the latter often regarded to be a more sophisticated approach. The lack of known-groups studies may stem partly from the lack of acceptance that "known group" designs actually reflect "malingerers".

Overall, there is clearly a substantial need for PVTs in GB with the growing claims culture but this is not reflected in the sophistication of the application of PVTs or the British research evidence base.

Germany

At the turn of the millennium, available German-language publications with the keyword malingering were a rarity and the topic of malingering was largely taboo, but an early review article (Heubrock & Petermann, 1995) apparently preceded later developments. Although two German-language tests (Eberl & Wilhelm, 2007; Heubrock & Petermann, 2000) were published and embedded measures were proposed by Schmidt-Atzert and colleagues (e.g., Schmidt-Atzert, Bühner, Rischen, & Warkentin, 2004) for the d2 Test of Attention, one of the most widespread German-language performance tests, controversy among (neuro)psychologists prevailed. German neuropsychologists were primarily identified with a clinical and therapeutic relationship to their patients and were little acquainted with the strictly impartial role a forensic expert has to assume. Still, in 2003, a prominent German neuropsychologist pronounced, in the context of assessing possible malingering, in a public neuropsychological newsgroup: "I am not inclined to witness and stand by –when neuropsychology moves back into military psychology." The therapeutic role to which psychologists are usually trained is implicitly linked with an attitude to believe what patients tell them. In the context of the ongoing debate about SVA, this is reflected by a statement of Noeker and Petermann (2011), who maintained that low scores on PVTs may erroneously suggest malingering in patients with functional and somatoform syndromes. Even if such patients produced an excessive array of symptoms, they nevertheless responded "in a subjectively truthful manner" (p. 450), and the inconsistency between different data should be understood by the expert as "the sensitive expression of major subjective distress and impairment" (p. 450).

Most empirical research focused on German language adaptations of PVTs (like the Amsterdam Short-Term Memory Test, the Medical Symptom Validity Test, and the Word Memory Test) as well as on the examination of embedded measures. A series of articles examining base rate estimates of negative response bias in forensic populations found numbers comparable to those known from North America (e.g., Stevens, Friedel, Mehren, & Merten, 2008). In accordance with more recent estimates from the United States, the prevalence of negative response bias approaches or exceeds the 50 % mark in some claimant populations (Merten, Thies, Schneider, & Stevens, 2009). Some research work found a larger interest among the scientific community, such as the first study to test a stand-alone symptom validity test in individuals with mental retardation (Brockhaus & Merten, 2004) and a study discussing and examining the *limits of effort testing* (Merten, Bossink, & Schmand, 2007), a topic which is very prominent in current developments. In the context of distorted self-report in the clinic (psychosomatic rehabilitation), a recent study by Göbber, Petermann, Piezga, and Kobelt (2012) found elevated scores on the Structured Inventory of Malingered Symptomatology (SIMS, Widows & Smith, 2005) in 33 % of German respondents and 50 % of respondents with migration background.

In the course of the last five or six years, SVA has increasingly been adapted by referral sources, including courts and state agencies, so much so that this is exerting pressure on hesitant or sceptical (neuro)psychologists and neurologists/psychiatrists to examine the genuineness of symptom report and the validity of test profiles with modern methods. Similarly to what is known from North America in the 1990s, referral agencies stipulate and stimulate the progress in the field. Today we have a somewhat paradoxical situation in Germany: PVT usage in medico-legal contexts appears to be fully accepted by *neurologists*, but the *neuropsychological* community continues to be divided. Many neuropsychologists are still hesitant to integrate modern approaches to SVA even into their forensic work. In *psychiatry*, there is an even deeper division. On the one hand, a growing number of psychiatrists appear to appreciate working closely with psychologists and include SVA results to make determinations about the genuineness of symptom report and the validity of profiles (e.g., Stevens, Fabra, & Merten, 2009). On the other hand, a group of influential German psychiatrists undertook to question the appropriateness of SVA in the context of forensic examinations in patients with mental disorders (e.g., Dressing, Foerster, Widder, Schneider, & Falkai, 2011). For a discussion of this deeply rooted controversy which has a profound impact on forensic psychiatric practice at the present time, see Schmidt, Lanquillon, and Ullmann (2011) and Merten and Merckelbach (2013b); an analysis of potential pitfalls for opponents of SVA was presented by Green and Merten (2013).

Remarkably, some sceptics or opponents to SVA appear to adopt the published English-language literature in a distorted way. To give a most striking example, the term *malingering* is wrongly understood as referring exclusively to symptom invention (or "pure" malingering), while other aspects of malingering are neglected. Thus, base rate estimates for malingered symptomatology are questioned on the basis of a wrong conceptualization (e.g., Dressing et al., 2011). However, a number of professional guidelines appear to be more circumspect, more open to include modern methods of SVA, and do not restrict psychiatric determinations to clinical decision making (e.g., Deutsche Rentenversicherung, 2012).

State of the art, recent developments, and outlook

While the European continent covers less than 7 % of the Earth's land area, it is quite diverse in terms of political and linguistic structure and comprises about 50 independent states with several dozens of official languages spoken by their native populations. Within the continent, neuropsychology certainly finds itself in diverse stages of development, and the role of forensic neuropsychology in the legal systems is probably as diverse (moreover, the states partly adhere to the inquisitorial, partly to the adversarial system, with different roles of the forensic expert within the legal system).

A similar diversity can be found in documented SVA practice and research activities as has been outlined above. In some countries, major developments have occurred to a degree that the gap between American and local SVA practice appears to have been narrowed while there is no documented proof of systematic SVA usage from most European countries. The recent statement of Sweet and Guidotti-Breting (2013) that there is no more controversy surrounding SVTs but "… this is merely a pseudo-controversy in that the number of neuropsychologists who hold this viewpoint now represent a very small minority" (p. 15) is certainly not true for Europe. A closer analysis of an ongoing harsh controversy in Switzerland and Germany was presented by Plohmann (2013) at the recent Third European Symposium on SVA.

Yet, professional guidelines for forensic assessment appear to include increasingly firm statements about SVT usage. However, as far as can be seen, Great Britain is the only country so far in which a special professional guideline dealing with SVA has been passed (McMillan et al., 2009).

Adoption of SVA in the majority of European countries is hampered by the diversity of national languages within the continent. While the development of original tests (like the ASTM in The Netherlands and TBFN and AST in Germany) is the exception, the bulk of SVA will have to rely on adaptions of international instruments (in particular of North American origin) as is the practice in neuropsychological assessment altogether. However, even adaptations of foreign language tests require proper validation. Moreover, the majority of practically working neuropsychologists in some major nations (like Germany, France, and Spain) are not used to resort to English language resources in their daily routines. To facilitate access to current developments, publications in their national languages appear to be mandatory (e.g., González-Ordi, Santamaría et al., 2012; Merten & Dettenborn, 2009). However, in a number of other countries (like The Netherlands and the Scandinavian nations), command of English language is a basic requirement for professionals.

Apart from the three major European symposia on Symptom Validity Assessment, a number of other events have marked the last few years. Workshops and national conference symposia on SVA were held in Austria, Germany, Great Britain, Switzerland, The Netherlands, and possibly in some more countries. In 2012, a workshop on the assessment of malingering and related response styles was held by Dr. Richard Rogers in Basel, Switzerland. Symposia on SVA were held at the World Congress of the International Brain Injury Association (IBIA), Edinburgh 2012, at two congresses of the International Academy of Law and Mental Health (IALMH), Berlin 2011 and Amsterdam 2013, at the annual meetings of the European Association of Psychology and Law (EAPL), Maastricht 2008, Gothenburg 2010, and Nicosia 2012, and at the joint meeting of the Federation of European Societies of Neuropsychology (ENS), Amsterdam 2010 and Berlin 2013.

With regard to SVA in Europe, it may be said that we have come a long way, but there is an even longer way before us. Whenever neuropsychologists enter the legal arena as independent experts, they have to deal with an elevated rate of invalid test profiles in claimants. SVA will help them to make valid determinations about possible response bias and to answer the legal questions with more confidence and more convincingly than if they rely on their clinical intuition. As a consequence, an appropriate and responsible use of these methods can –and will– enhance the competence and the esteem of neuropsychologists in the legal arena. In clinical and rehabilitation contexts SVA can play a significant role in differential diagnosis and help clarifying the nature of certain symptom constellations as well as factors contributing to the course of reported symptoms. In these contexts, however, the way to go is even longer.

Conflicts of interest

Dr. Merten, Prof. Schmand, Prof. González-Ordi, and Santamaría have participated in the adaptations of numerous validity tests to German, Dutch and Spanish respectively. Prof. Schmand is the author of the Amsterdam Short-Term Memory Test. No conflicts of interest were declared by the other authors.

References

Alcott, D. (2006). Post-qualification training day on symptom validity testing attracts capacity audience. *Newsletter, The British Psychological Society Division of Neuropsychology, 5,* 1-2.

Anderson, S. (2010). Conference report: The First European Symposium on Symptom Validity Assessment (May 8 – 9 2009). *Zeitschrift für Neuropsychologie, 21*, 65.

Arce, R., Fariña, F., Carballal, A., & Novo, M. (2009). Creación y validación de un protocolo de evaluación forense de las secuelas psicológicas de la violencia de género [Development and validation of a protocol for forensic evaluation of psychological sequelae of sexual violence]. *Psicothema. 21,* 241–247.

Arce, R., & Fariña, F. (2005). Peritación psicológica de la credibilidad del testimonio, la huella psíquica y la simulación: El sistema de evaluación global (SEG) [Psychological expert opinion on the credibility of witness accounts, on psychological injury and malingering: The Global Evaluation System, GES]. *Papeles del Psicólogo, 26,* 59–77.

Association of British Insurers (2012). *No hiding place insurance fraud exposed*. Retrieved from https://www.abi.org.uk/~/media/Files/Documents/Publications/Public/Migrated/Fraud/ABI%20no%20hiding%20place%20-%20insurance%20fraud%20exposed.ashx

Benton, A. L. (1945). A visual retention test for clinical use. *Archives of Neurology and Psychiatry, 64,* 212-216.

Benton, A. L., & Spreen, O. (1961). Zur Simulation intellektueller Leistungsdefekte im Benton-Test [About malingered deficits of intelligence performance in the Visual Retention Test]. *Psychologische Beiträge, 7,* 147-150.

Bianchini, K. J., Greve, K. W., & Glynn, G. (2005). On the diagnosis of malingered pain-related disability: Lessons from cognitive malingering research. *The Spine Journal,5,* 404-417.

Brockhaus, R., & Merten, T. (2004). Neuropsychologische Diagnostik suboptimalen Leistungsverhaltens mit dem Word Memory Test [Neuropsychological assessment of suboptimal performance: The Word Memory Test]. *Nervenarzt, 75,* 882-887.

Capilla-Ramírez, P., & González-Ordi, H. (2009). Protocolo para la detección de la simulación del dolor en la práctica clínica: estudio de casos [Clinical protocol for the detection of malingered pain: Case studies]. *Trauma, 20,* 255–263.

Carone, D. A., & Bush, S. S. (2013). Introduction: Historical perspectives on mild traumatic brain injury, symptom validity assessment, and malingering. In D. A. Carone & S. S. Bush. (Eds.), *Mild traumatic brain injury: Symptom validity assessment and malingering* (pp. 1-29). New York: Springer.

Castiel, M., Alderman, N., Jenkins, K., Knight, C., & Burgess, P. (2012). Use of the Multiple Errands Test–Simplified Version in the assessment of suboptimal effort. *Neuropsychological Rehabilitation, 22*, 734-751.

Chafetz, M., & Underhill, J. (in press). Estimated costs of malingered disability. *Archives of Clinical Neuropsychology*.

Dandachi-FitzGerald, B., Ponds, R., & Merten, T. (in press). Symptom validity and neuropsychological assessment: A survey of practices and beliefs in six European countries. *Archives of Clinical Neuropsychology*. doi:10.1093/arclin/act073

Dandachi-FitzGerald, B., Ponds, R. W. H. M., Peters, M. J. V., & Merckelbach, H. (2011). Cognitive underperformance and symptom over-reporting in a mixed psychiatric sample. *The Clinical Neuropsychologist, 25*, 812-828.

Department for Work & Pensions (2013). *Simplifying the welfare system and making sure work pays.* London: The Stationary Office.

Deutsche Rentenversicherung (2012). *Leitlinien zur sozialmedizinischen Beurteilung. Sozialmedizinische Beurteilung bei psychischen und Verhaltensstörungen* [Professional guidelines of the German Pension Fund for independent medical examinations: Mental and behavioral disorders]. Berlin: Author.

Dressing, H., Foerster, K., Widder, B., Schneider, F., & Falkai, P. (2011). Zur Anwendung von Beschwerdenvalidierungstests in der psychiatrischen Begutachtung [On the application of symptom validity tests in psychiatric forensic examination]. *Nervenarzt, 82,* 388-390.

Eberl, A., & Wilhelm, H. (2007). *Aggravations- und Simulationstest AST 4.0* [Exaggeration and Malingering Test] (2nd edition). Essen: Mnemo-Verlag.

Erichsen, J. E. (1867*). On railway and other injuries of the nervous system.* Philadelphia, PA: Henry C. Lea.

Fisher, H. L., & Rose, D. (2005). Comparison of the effectiveness of two versions of the Rey memory test in discriminating between actual and simulated memory impairment, with and without the addition of a standard memory test. *Journal of Clinical and Experimental Neuropsychology, 27*, 840-858.

Frederick, R. I. (2002). A review of Rey's strategies for detecting malingered neuropsychological impairment. *Journal of Forensic Neuropsychology, 2,* 1-25.

Frei, M. (2004, July 29). Wie sich Simulanten verraten [How malingerers reveal themselves]. *Tages-Anzeiger*, p. 26.

García-Silgo, M., & Robles-Sánchez, J. I. (2010). Estudio de las bajas temporales para el servicio por motivos psicológicos: mejoras en la detección de psicopatología y simulación de trastorno mental en las Fuerzas Armadas [Malingering in psychological sick leaves in the military]. *Sanidad Militar: Revista de Sanidad de las Fuerzas Armadas de España, 66,* 154–162.

Giger, P., Merten, T., Merckelbach, H., & Oswald, M. (2010). Detection of feigned crime-related amnesia: A multi-method approach. *Journal of Forensic Psychology Practice, 10,* 140-163.

Gill, D., Green, P., Flaro, L., & Pucci, T. (2007). The role of effort testing in independent medical examinations. *Medico-Legal Journal, 75,* 64-71.

Göbber, J., Petermann, F., Piegza, M., & Kobelt, A. (2012). Beschwerdenvalidierung bei Rehabilitanden mit Migrationshintergrund in der Psychosomatik [Symptom validation in patients with migration background in psychosomatic medicine]. *Rehabilitation, 51,* 356-364.

González-Ordi, H., & Gancedo-Rojí, M. (1999). Evaluación de la simulación de respuesta: diagnóstico, técnicas y procedimientos [Malingering: Assessment, techniques, and procedures]. In A. Ruano, J. M. Muñoz-Céspedes, & C. Cid-Rojo (Coords.), *Psicología de la Rehabilitación* (pp. 405-441). Madrid: Fundación MAPFRE-Medicina.

González-Ordi, H., Santamaría, P., & Fernández-Marín, P. (2010). Precisión predictiva del Inventario Estructurado de Simulación de Síntomas – SIMS en el contexto médicolegal [Predictive accuracy of the Structured Inventory of Malingered Symptomatology—SIMS in the medicolegal setting]. *Edupsykhé, 9,* 3-22.

González-Ordi, H., Capilla-Ramírez, P., Santamaría, P., & Casado-Morales, M. I. (2012). Abordaje multidisciplinar para la detección de la simulación en lumbalgia crónica. [A multidisciplinary approach to the detection of malingering in chronic lower back pain]. *Trauma Fundación MAPFRE, 23*, 145–154.

González-Ordi, H., & Santamaría, P. (2009). *Adaptación española del Inventario Estructurado de Simulación de Síntomas* [Spanish adaptation of the Structured Inventory of Malingered Symptomatology]. Madrid: TEA Ediciones.

González-Ordi, H., Santamaría, P., & Capilla-Ramírez, P. (Eds.). (2012*). Estrategias de detección de la simulación: Un manual clínico multidisciplinar* [Malingering detection strategies: A multidisciplinary clinical manual]. Madrid: TEA Ediciones.

Gorissen, M., Sanz, J. C., & Schmand, B. (2005). Effort and cognition in schizophrenia patients. *Schizophrenia Research, 78*, 199-208.

Green, P. (2003). *Green's Word Memory Test: User's manual*. Edmonton, Canada: Green's Publishing.

Green, P. (2004). *Green's Medical Symptom Validity Test (MSVT) for Microsoft Windows: User's Manual*. Edmonton, Canada: Green's Publishing.

Green, P., & Merten, T. (2013). Noncredible explanations of noncredible performance on symptom validity tests. In D. A. Carone & S. S. Bush. (Eds.), *Mild traumatic brain injury: Symptom validity assessment and malingering* (pp. 73-100). New York: Springer.

Greve, K. W., Bianchini, K. J., & Brewer, S. T. (2013). The assessment of performance and self-report validity in persons claiming pain-related disability. *The Clinical Neuropsychologist, 27,* 108–137.

Grose, F. (1811*). Dictionary of the vulgar tongue. A Dictionary of buckish slang, university wit, and pickpocket eloquence.* Retrieved from http://www.freeinfosociety.com/media/pdf/3934.pdf

Gudjonsson, G. H., & Shackleton, H. (1986). The pattern of scores on Raven's Matrices during 'faking bad' and 'non-faking' performance. *British Journal of Clinical Psychology, 25,* 35-41.

Hacker, V. L., & Jones, C. (2009). Detecting feigned impairment with the word list recognition of the Wechsler Memory Scale. *Brain Injury, 23,* 243-249.

Hall, V. L., Worthington, A. & Venables, K. (in press). A UK pilot study: The specificity of the Word Memory Test effort sub-tests in acute minimal to mild head injury. *Journal of Neuropsychology.*

Halligan, P. W., Bass, C., & Oakley, D. A. (Eds.) (2003). *Malingering and illness deception.* Oxford: Oxford University Press.

Hanley, J. R., Baker, G. A., & Ledson, S. (1999). Detecting the faking of amnesia: A comparison of the effectiveness of three different techniques for distinguishing simulators from patients with amnesia. *Journal of Clinical and Experimental Neuropsychology,21*, 59-69.

Hartman, D. E. (2002). The unexamined lie is a lie worth fibbing. Neuropsychological malingering and the Word Memory Test. *Archives of Clinical Neuropsychology, 17,* 709-714.

Heaton, R. K., Smith, H. H., Lehman, R. A. W., & Vogt, A. T. (1978). Prospects for faking believable deficits on neuropsychological testing. *Journal of Consulting and Clinical Psychology, 46*, 892-900.

Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. J., Millis, S. R., & Conference Participants (2009). American Academy of Clinical Neuropsychology consensus conference statement on the neuropsychological assessment of effort, response bias, and malingering. *The Clinical Neuropsychologist, 23,* 1093-1129.

Heubrock, D. (1995). Neuropsychologische Diagnostik bei Simulationsverdacht - ein Überblick über Forschungsergebnisse und Untersuchungsmethoden [Neuropsychological assessment of suspected malingering: Research results and evaluation techniques]. *Diagnostica, 41,* 303-321.

Heubrock, D., & Petermann, F. (1998). Neuropsychological assessment of suspected malingering: Research results, evaluation techniques, and further directions of research and application. *European Journal of Psychological Assessment, 14,* 211-225.

Heubrock, D., & Petermann, F. (2000). *Testbatterie zur Forensischen Neuropsychologie, TBFN. Testmanual. Neuropsychologische Diagnostik bei Simulationsverdacht* [Forensic Neuropsychological Test Battery. Manual]. Frankfurt: Swets Test Services.

Hofstee, W. K. B. (2005). De psycholoog als detective? Kanttekeningen bij malingering- en integriteitstests [The psychologist as a detective? Comments on malingering and integrity tests]. *De Psycholoog, 40*, 670-674.

Jenkins, K. G., Kapur, N., & Kopelman, M. D. (2009). Retrograde amnesia and malingering. *Current Opinion in Neurology, 22,* 601-605.

Jiménez-Gómez, F., & Sánchez-Crespo, G. (2004). *La falsificación de las técnicas psicométricas: un estudio con el MMPI-2* [The falsification of psychometric techniques: A study with the MMPI-2]. Salamanca: Universidad de Salamanca.

Kapur, N. (1994). The Coin-in-the-Hand Test: A new "bedside" test for the detection of malingering in patients with suspected memory disorder. *Journal of Neurology, Neurosurgery and Psychiatry, 57,* 385-386.

Kelly, P. J., Baker, G. A., Broek, M. D., Jackson, H., & Humphries, G. (2005). The detection of malingering in memory performance: The sensitivity and specificity of four measures in a UK population. *British Journal of Clinical Psychology, 44,* 333-341.

Kemp, S., Coughlan, A. K., Rowbottom, C., Wilkinson, K., Teggart, V., & Baker, G. (2008). The base rate of effort test failure in patients with medically unexplained symptoms. *Journal of Psychosomatic Research, 65,* 319-325.

Larrabee, G. J. (2005). Assessment of malingering. In G. J. Larrabee (Ed.), *Forensic neuropsychology: A scientific approach* (pp. 115–158). New York: Oxford University Press.

Larrabee, G. J. (2012). Performance validity and symptom validity in neuropsychological assessment. *Journal of the International Neuropsychological Society, 18,* 625-631.

Larrabee, G. J., Greiffenstein, M. F., Greve, K. W., & Bianchini, K. J. (2007). Refining diagnostic criteria for malingering. In G. J. Larrabee (Ed.), *Assessment of* malingered neuropsychological deficits (pp. 334-371). New York: Oxford University Press.

Lemos-Giráldez, S. (2005). Simulación, engaño y mentira [Malingering, deception and lies]. *Papeles del Psicólogo, 26,* 57–58.

Loewer, H. D., & Ulrich, K. (1971). Eine Alternativ-Wahl-Form des Benton-Testes zur besseren Erfassung von Aggravation und Simulation [A forced-choice version of the Visual Retention Test for improved assessment of malingering]. In E. Duhm (Hrsg.), *Praxis der klinischen Psychologie*, Band 2 (pp. 63-75). Göttingen: Hogrefe.

Martins, M., & Martins, I. P. (2010). Memory malingering: Evaluating WMT criteria. *Applied Neuropsychology, 17,* 177-182.

Massey, R. (2013, May 27). Half of whiplash claims are bogus: Faked injuries driving a £1billion-a-year racket. *The Daily Mail*. Retrieved from http://www.dailymail.co.uk/news/article-2331858/Half-whiplash-claims-bogus-Faked-injuries-driving-1billion-year-racket.html

Masip, J., Garrido, E., & Herrero, C. (2009). Heuristic versus systematic processing of information in detecting deception: Questioning the truth bias. *Psychological Reports, 105,* 11–36.

McCarter, R. J., Walton, N. H., Brooks, D. N., & Powell, G. E. (2009). Effort testing in contemporary UK neuropsychological practice. *The Clinical Neuropsychologist, 23,* 1050-1066.

McMillan, T. M., Anderson, S., Baker, G., Berger, M., Powell, G. E., & Knight, R. (2009) Assessment of effort in clinical testing of cognitive functioning for adults. Leicester, UK: British Psychological Society.

Merckelbach, H., Dandachi-FitzGerald, B., Van Mulken, P., Ponds, R., & Niesten, I. (in press). Exaggerating psychopathology produces residual effects that are resistant to corrective feedback: An experimental demonstration. *Applied Neuropsychology: Adult*.

Merckelbach, H., Jelicic, M., Pieters, M. (2011). The residual effect of feigning: how initial feigning may evolve into a less conscious form of symptom reporting. *Journal of Clinical and Experimental Neuropsychology, 33*, 131-139.

Merckelbach, H., & Smith, G. P. (2003). Diagnostic accuracy of the Structured Inventory of Malingered Symptomatology (SIMS) in detecting malingering. *Archives of Clinical Neuropsychology, 18*, 145-152.

Merten, T., Bossink, L., & Schmand, B. (2007). On the limits of effort testing: Symptom validity tests and severity of neurocognitive symptoms in nonlitigant patients. *Journal of Clinical and Experimental Neuropsychology, 29,* 308-318.

Merten, T., & Dettenborn, H. (Eds.) (2009). *Diagnostik der Beschwerdenvalidität* [Symptom validity assessment]. Berlin: Deutscher Psychologen Verlag.

Merten, T., & Merckelbach, H. (2013a). Forced-choice tests as experiments in the differential diagnosis of intentional symptom distortion. *Journal of Experimental Psychopathology, 4,* 20-37.

Merten, T., & Merckelbach, H. (2013b). Symptom validity testing in somatoform and dissociative disorders: A critical review. *Psychological Injury and Law, 6,* 122-137.

Merten, T., Thies, E., Schneider, K., & Stevens, A. (2009). Symptom validity testing in claimants with alleged posttraumatic stress disorder: Comparing the Morel Emotional Numbing Test, the Structured Inventory of Malingered Symptomatology, and the Word Memory Test. *Psychological Injury and the Law, 2,* 284-293.

Millis, S. R. (1994). Assessment of motivation and memory with the Recognition Memory Test after financially compensable mild head injury. *Journal of Clinical Psychology, 50,* 601-605.

Ministry of Justice (2013). *Written Evidence from the Misistry of Justice (WL 55).* Retrieved from http://www.publications.parliament.uk/pa/cm201314/cmselect/cmtran/writev/whiplash/wl55-56.pdf

Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base Rates of Malingering and Symptom Exaggeration. *Journal of Clinical and Experimental Neuropsychology, 24,* 1094-1102.

Moss, A., Jones, C., Fokias, D., & Quinn, D. A. (2003) The mediating effects of effort upon the relationship between head injury severity and cognitive functioning. *Brain Injury, 17,* 377-387.

Noeker, M., & Petermann, F. (2011). Simulation neurologischer versus psychischer Beschwerden. Notwendigkeit unterschiedlicher Validierungsstrategien [Malingering of neurological vs. mental complaints: Necessity of different validation strategies]. *Psychotherapeut, 56,* 449-454.

Ortiz-Tallo, M., Santamaría, P., Cardenal, V., & Sánchez, P. (2011). *Adaptación española del PAI* [Spanish adaptation of the PAI]. Madrid: TEA Ediciones.

Pankratz, L. (1983). A new technique for the assessment and modification of feigned memory deficit. *Perceptual and Motor Skills, 57,* 367-372

Pankratz, L., & Paar, G. (1988). Test zur Symptomvalidität zur Einschätzung funktioneller Symptome [A symptom validity test for the evaluation of functional symptoms]. *Zeitschrift für Klinische Psychologie, Psychopathologie und Psychotherapie, 36,* 130-137.

Plohmann, A. M. (2011). Zweite Europäisches Symposium zur Diagnostik der Beschwerdenvalidität. Tagungsbericht [Second European Symposium on Symptom Validity Assessment: Conference report]. *Zeitschrift für Neuropsychologie, 22,* 303-305.

Plohmann, A. (2013, June). *Common arguments against SV assessment: The psychiatry debate in Germany and Switzerland.* Paper presented at the Third European Symposium on Symptom Validity Assessment. Würzburg, Germany.

Plohmann, A., & Merten, T. (2013). The third European Symposium on Symptom Validity Assessment. Facts and controversies. *Clínica y Salud, 25,* 197-203.

Ponds, R., de Lugt, M., Verhey, F., & Jolles, J. (1995). Malingering bij neuropsychologisch onderzoek [Malingering in neuropsychological assessment]. *De Psycholoog, 30,* 357-361.

Rey, A. (1941). L'examen psychologique dans les cas d'encéphalopathie traumatique [The psychological examination in cases of traumatic encephalopathy]. *Archives de Psychologie, 28,* 286-340.

Rey, A. (1958). *L'examen clinique en psychologie* [The clinical examination in psychology]. Paris: Presses Universitaires de France.

Rienstra, A., Groot, P. F. C., Spaan, P. E. J., Majoie, C. B. L. M., Nederveen, A. J., Walstra, G. J. M., … Schmand, B. (2013). Symptom validity testing in memory clinics: hippocampal-memory associations and relevance for diagnosing mild cognitive impairment. *Journal of Clinical and Experimental Neuropsychology, 35*, 59-70.

Rienstra, A., Klein Twennaar, M., & Schmand, B. (2013). Neuropsychological characterization of patients with the WMT dementia profile. *Archives of Clinical Neuropsychology, 28*, 463-475.

Sánchez-Crespo, G., Jiménez-Gómez, F., Ampudia-Rueda, A., & Merino-Barragán, V. (2012). In search of a fast screening method for detecting the malingering of cognitive impairment. *European Journal of Psychology Applied to Legal Context, 4,* 135–158.

Santamaría, P. (2009). *Adaptación española del MMPI-2-RF* [Spanish adaptation of the MMPI-2-RF]. Madrid: TEA Ediciones.

Schiemann, S. (2003). Entwicklung und Erprobung einer Testbatterie zur Diagnostik bei Simulationsverdacht [Development and evaluation of a test battery for the detection of malingering]. *Psychology Science, 45,* 80-100.

Schmand, B., de Sterke, S., & Lindeboom, J. (1999). *Amsterdamse Korte Termijn Geheugen test* [Amsterdam Short-Term Memory Test]. Lisse, NL: Swets & Zeitlinger.

Schmand, B., & Lindeboom, J. (2005). Amsterdam Short-Term Memory Test. Manual. Leiden, The Netherlands: PITS.

Schmand, B., Lindeboom, J., Schagen, S., Heijt, R., Koene, T., & Hamburger, H. L. (1998). Cognitive complaints in patients after whiplash injury: The impact of malingering. *Journal of Neurology, Neurosurgery, and Psychiatry, 64,* 339-343.

Schmand, B., & Ponds, R. (1997). Malingeren: simuleren en aggraveren [Malingering: simulation and aggravation]. In: B. Deelman, P. Eling, E. de Haan, A. Jennekens-Schinkel, E. van Zomeren (Eds.), *Klinische Neuropsychology* [Clinical Neuropsychology] (pp. 426-436). Amsterdam: Boom.

Schmidt, T., Lanquillon, S., & Ullmann, U. (2011). Kontroverse zu Beschwerdenvalidierungsverfahren bei der Begutachtung psychischer Störung [Controversy of symptom validity assessment in forensic evaluation of mental disorders]. *Forensische Psychiatrie Psychologie Kriminologie, 5,* 177-183.

Schmidt-Atzert, L., Bühner, M., Rischen, S., & Warkentin, V. (2004). Erkennen von Simulation und Dissimulation im Test d2 [Detection of malingering and dissimulation in d2 test]. *Diagnostica, 50,* 124-133.

Schroeder, R. W., Peck, C. P., Buddin, W. H., Heinrichs, R. J., & Baade, L. E. (2012). The Coin-in-the-Hand Test and dementia: More evidence for a screening test for neurocognitive symptom exaggeration. *Cognitive and Behavioral Neurology, 25,* 139-143.

Simões, M., Sousa, L., Duarte, P., Firmino, H., Pinho, M. S., Gaspar, N., … França, S. (2010). Avaliação da simulação ou esforço insuficiente com o Rey 15-Item Memory Test (15-IMT): Estudos de validação em grupos de adultos idosos [Validation of simulation or insufficient effort with the Rey 15-IMT: Validation studies in aged adult groups]. *Análise Psicológica, 28,* 209–226.

Singhal, A., Green, P., Ashaye, K., Shankar, K., & Gill, D. (2009). High specificity of the Medical Symptom Validity Test in patients with very severe memory impairment. *Archives of Clinical Neuropsychology, 24,* 721-728.

Slick, D. J., Sherman, E. M., & Iverson, G. L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research.*The Clinical Neuropsychologist, 13,* 545-561.

Spreen, O. (1963). *MMPI-Saarbrücken. Handbuch zur deutschen Ausgabe des MMPI von S. R. Hathaway und J. C. McKinley* [MMPI Saarbrücken: Manual of the German edition of the MMPI by S. R. Hathaway and J. C. McKinley]. Bern: Huber.

Squires, N. (2013, August 01). Greece tries to crack down on fraud as mayor of Zakynthos faces revolt. *The Telegraph.* Retrieved from http://www.telegraph.co.uk/news/worldnews/europe/greece.

Stevens, A., Fabra, M., & Merten, T. (2009). Anleitung für die Erstellung psychiatrischer Gutachten [Directions for forensic psychiatric examinations]. *Der medizinische Sachverständige, 105,* 100-106.

Stevens, A., Friedel, E., Mehren, G., & Merten, T. (2008). Malingering and uncooperativeness in psychiatric and psychological assessment: Prevalence and effects in a German sample of claimants. *Psychiatry Research, 157,* 191-200.

Sweet, J. J., & Guidotti-Breting, L. M. (2013). Symptom validity test research: Status and clinical implications. *Journal of Experimental Psychopathology, 4,* 6-19.

Thomann, K. D., & Rauschmann, M. A. (2003). Von der „railway spine" zum Schleudertrauma: Geschichte und aktuelle Bedeutung seelischer Störungen nach entschädigungspflichtigen Ereignissen [From "railway spine" to whiplash injury]. *Zeitschrift für die gesamte Versicherungswissenschaft, 92,* 533-577.

Tombaugh, T. N. (1996). *Test of Memory Malingering (TOMM).* North Tonawanda, NY: Multi-Health Systems.

Vallabhajosula, B., & van Gorp, W. G. (2001). Post-Daubert admissibility of scientific evidence on malingering of cognitive deficits. *Journal of the American Academy of Psychiatry Law, 29*, 207–215.

Van der Werf, S. P., Prins, J. B., Jonge, P. J., van der Meer, J. W., & Bleijenberg, G. (2000). Abnormal neuropsychological findings are not necessarily a sign of cerebral impairment: a matched comparison between chronic fatigue syndrome and multiple sclerosis. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology, 13*, 199-203.

Van Zomeren, E., & Saan, R. (1997). Whiplash. In B. Deelman, P. Eling, E. de Haan, A. Jennekens-Schinkel, & E. van Zomeren (Eds.), *Klinische Neuropsychology* [Clinical Neuropsychology] (pp. 290-298). Amsterdam: Boom.

Vilar-López, R., & Aliaga, A. (2009). Evaluación de la simulación de problemas cognitivos [Assessment of cognitive problems malingering]. In A. Jarne, T. Villaseñor, & A. Aliaga (Eds.), *Manual de neuropsicología forense*. Barcelona: Paidós.

Vilar-López, R., Gómez-Río, M., Caracuel-Romero, A., Llamas-Elvira, J., & Pérez-García, M. (2008). Use of specific malingering measures in a Spanish sample. *Journal of Clinical and Experimental Neuropsychology, 30,* 710-722.

Vilar-López, R., Pérez-García, M., & Puente, A. (2011). *Adaptación española del Test de Simulación de Problemas de Memoria TOMM* [Spanish adaptation of the Test of Memory Malingering]. Madrid: TEA Ediciones.

Warrington, E. K. (1984). *Recognition Memory Test.* Windsor, England: NFER-Nelson.

Widows, M. R., & Smith, G. P. (2005). *SIMS-Structured Inventory of Malingered Symptomatology.* Professional manual. Lutz, FL: Psychological Assessment Resources.

Wogar, M. A., van den Broek, M. V. D., Bradshaw, C. M., & Szabadi, E. (1998). A new performance-curve method for the detection of simulated cognitive impairment. *British Journal of Clinical Psychology, 37,* 327-339.

Zaldivar-Basurto, F., García-Montes, J. M., López-Rios, F., Molina-Moreno, A., & Santiago, E. (2007). *Adaptación y validación al castellano de la Evaluación Forense de Miller del Test de Síntomas (M-FAST)* [Spanish adaptation and validation of the Miller-Forensic Assessment of Symptoms Test]. Paper presented at the 3rd Congreso de Psicología Jurídica y Forense. Oviedo: Ediciones de la Universidad de Oviedo.