



# Symptom and Performance Validity Assessment in European Countries: an Update

Thomas Merten<sup>1</sup> · Brechje Dandachi-FitzGerald<sup>2</sup> · Vicki Hall<sup>3</sup> · Thomas Bodner<sup>4</sup> · Luciano Giromini<sup>5</sup> · Johann Lehrner<sup>6</sup> · Héctor González-Ordi<sup>7</sup> · Pablo Santamaría<sup>8</sup> · Ben Schmand<sup>9</sup> · Giuseppe Di Stefano<sup>10</sup>

Received: 31 August 2021 / Accepted: 13 November 2021 / Published online: 24 November 2021  
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

## Abstract

In 2013, a special issue of the Spanish journal *Clínica y Salud* published a review on symptom and performance validity assessment in European countries (Merten et al. in *Clínica y Salud*, 24(3), 129–138, 2013). At that time, developments were judged to be in their infancy in many countries, with major publication activities stemming from only four countries: Spain, The Netherlands, Great Britain, and Germany. As an introduction to a special issue of *Psychological Injury and Law*, this is an updated report of developments during the last 10 years. In that period of time, research activities have reached a level where it is difficult to follow all developments; some validity measures were newly developed, others were adapted for European languages, and validity assessment has found a much stronger place in real-world evaluation contexts. Next to an update from the four nations mentioned above, reports are now given from Austria, Italy, and Switzerland, too.

**Keywords** Validity · Symptom validity test · Performance validity test · Malingering · Psychological assessment

## Symptom Validity Assessment in Europe About 10 Years Ago

It is almost 10 years ago that Prof. González-Ordi from Complutense University, Madrid, asked the first author to contribute to a special issue on symptom validity

assessment (SVA) to be published in the Spanish Elsevier journal *Clínica y Salud*. He did this at the third European Symposium on SVA held in Wuerzburg, Germany, in 2013, organized by the International Academy of Applied Neuropsychology (led by Gerhard Müller and Herbert König). The result of that request was a historical sketch on symptom validity assessment in Europe, beginning with the pioneering work of Rey (1941) and comprising state of the art reports contributed by colleagues from those four European nations that were most visible in SVA research at that time: the Netherlands, Spain, Great Britain, and Germany.

The current article is an attempt to update that text (Merten et al., 2013) without repeating the information given there (the text is available free of charge from the publisher; see References). In a fast-developing field of research and assessment practice like SVA, the time frame of a decade is likely to bring about significant changes. The term symptom validity was historically developed, in the 1970s (e.g., Pankratz, 1979) but by 2013, it was no longer used as a superordinate concept. Larrabee (2012) had proposed to (verbally) differentiate between symptom validity tests (SVTs, from now on relating mostly to self-report validity measures, today also comprising interview methods) and performance validity tests (PVTs, relating to cognitive validity measures).

✉ Thomas Merten  
thomas.merten@vivantes.de

- <sup>1</sup> Department of Neurology, Vivantes Klinikum Im Friedrichshain, Berlin, Germany
- <sup>2</sup> Department of Clinical Psychological Science, Maastricht University, Maastricht, The Netherlands
- <sup>3</sup> Neuromindworks Director, Sutton Coldfield, Great Britain
- <sup>4</sup> Department of Neurology, Medical University of Innsbruck, Innsbruck, Austria
- <sup>5</sup> Department of Psychology, University of Turin, Turin, Italy
- <sup>6</sup> Department of Neurology, Medical University of Vienna, Vienna, Austria
- <sup>7</sup> Faculty of Psychology, Complutense University of Madrid, Madrid, Spain
- <sup>8</sup> TEA Ediciones, Madrid, Spain
- <sup>9</sup> Brain & Cognition, University of Amsterdam, Amsterdam, The Netherlands
- <sup>10</sup> IB-Bern GmbH, Berne, Switzerland

It was also at that time that the findings of a European survey on symptom validity testing (in the historical, superordinate sense of the term) were published (Dandachi-FitzGerald et al., 2013), following a previous survey from Great Britain (McCarter et al., 2009). Efforts to motivate as many national neuropsychological societies as possible to participate resulted in responses given by neuropsychologists from Denmark, Finland, Germany, Italy, the Netherlands, and Norway. Note should be taken that half of the national neuropsychological societies that had been contacted and asked to participate in the survey, either did not respond to repeated requests or signaled that, in their opinion, the study was not feasible in their countries.

About 10 years ago, there was considerable resistance against SVA even among forensically working neuropsychologists, on the background of an irrational conviction that a professional could easily tell apart genuine from manipulated symptom presentations without having to resort to special means and methods. Even more resistance was notable among many psychiatrists who apparently felt that their traditional intuitive approach of relying on subjective symptom report by patients (without thoroughly investigating its authenticity and possible significant response distortions) was threatened by the introduction of empirically based methods and a data-driven approach. These were methods many psychiatrists did not use and did not understand. This was clearly visible in both Germany and Switzerland (e.g., Dressing et al., 2011). For a more detailed account on forms of resistance against SVA as it had emerged in the 1990s and early 2000s, see Green and Merten (2013).

The whole dispute bears some resemblance with the old controversy of clinical versus statistical predictions (Meehl, 1954). The basic problem here is which kind of data is superior for arriving at valid diagnoses and prognoses. We have learned that, under some circumstances, statistical predictions are not automatically superior to clinical judgment. This appears to apply to some contexts, such as the classification of seizure types (e.g., Fargo et al., 2008), and judgments from persons with special clinical expertise (Ægisdóttir et al., 2006). The same may also apply to some forensic decision making contexts; a combination of statistical and clinical data may, in fact, turn out to be superior to either of them in isolation. In this vein, the new multidimensional malingering criteria for neuropsychological assessment (Sherman et al., 2020) also comprise specifiers for the clinical presentation of malingering. However, the practitioner should always bear in mind that human judges often overestimate their abilities (Kahneman et al., 2021).

In Britain, the use of the term “malingering” in court-ordered forensic expert reports continued to be largely taboo (see more detailed report below). The further development of the whiplash crisis in Britain (as described in the Merten et al., 2013, report) and the public perception of fraudulent symptom

claims after motor vehicle accidents, were dealt with in a series of articles by Cartwright and his co-workers (e.g., Cartwright & Roach, 2015; Cartwright et al., 2019).

Around the turn of the millennium, there was an apparent delay in SVA of about 10 years in Europe, as compared to Northern American developments. For many psychological and medical professionals, even for some forensically working experts, the topic of feigned symptom presentations was largely taboo. The conclusion of the Merten et al. (2013) review was that parts of Central and Western Europe were about to reduce the delay in SVA research and practice significantly while in other parts of the continent (large parts, to be sure), no major published research was detectable. Yet, available European estimates of invalid responding and uncooperativeness in civil and social-law forensic contexts pointed at base rates similar to those obtained in North America (e.g., Allcott et al., 2014; Merten et al., 2020; Plohmman & Hurter, 2017).

It was the former member of the Executive Committee of the British Psychological Society, Division of Neuropsychology, Dr. Stuart Anderson, who formulated the idea of organizing and convening a European symposium on SVA (Anderson, 2010). This first meeting in Wuerzburg, Germany, was felt to be such an extraordinary success that subsequent symposia were held every other year. As a result, participants could keep track of the latest developments in this fast-evolving field and, most of all, meet and hear a selection of the most important experts; the list of contributors and keynote speakers reads like a *Who is Who* in symptom and performance validity research. Six conferences were held in Germany, The Netherlands, Great Britain, and Switzerland before the 2-year rhythm was unexpectedly interrupted by the COVID-19 pandemic.

The program list comprises invited contributors from at least 12 countries. Poster sessions were held at each conference, and poster blitz presentations were perceived to be a powerful way of alerting the audience to new developments, newly conceived tests, unpublished new studies, research projects in their planning phase, single-case studies, etc.

## European Developments in Symptom Validity Assessment

The following reports embark on describing the state of the art in a number of European countries, those that were most visible in the SVA literature. For Great Britain, the Netherlands, Germany, and Spain, previous reports can be found in Merten et al. (2013), so it is updated from the earlier accounts only that were included here. Full reports were requested from all other contributing countries, that is, Austria, Italy, and Switzerland. Despite multiple efforts, no information could be obtained from a number of other countries, including France and the Scandinavian countries.

## Germany

The awareness of possible significant response distortions in forensic assessments has grown further in the past 10 years, with an ongoing debate among psychiatrists about the use of SVTs and PVTs in patients with claimed mental disorders. Despite this debate, many psychiatrists began to use self-report validity measures, in particular the Structured Inventory of Malingered Symptomatology (SIMS; Smith & Burger, 1997), but the problem of correctly handling and interpreting results of this questionnaire and other instruments was visible in many expert reports. As with psychological testing in general, many non-psychologists continue to underestimate both the complexity of psychological assessment (symptom and performance validity testing included, of course) and the sound qualification needed to correctly use tests and interpret their results.

The discovery and publication of huge fraud networks targeting social security and social welfare schemes in Germany (Deutsche Rentenversicherung, 2017; Hoffmann, 2019) should have enabled even the most conservative of all critics of symptom validity research to correct their persistent belief that malingering and fraudulent disability claims were rare phenomena relevant for American, but not for European or German social realities. But old generals and irrational convictions never die, so resistance against SVA will most probably only fade away with time.

Most visible among newly developed validity measures and published research were the Beschwerdvalidierungstest (BEVA; Walter et al., 2016) and the Self-Report Symptom Inventory (SRSI; Merten et al., 2016). At about the same time, German adaptations of the Structured Interview of Reported Symptoms–2 (SIRS-2; Schmidt et al., 2019) and the Inventory of Problems–29 (IOP-29; Viglione & Giromini, 2020) were tested and made available to German-language users. Among PVTs, the Groningen Effort Test (GET; Fuermaier et al., 2017) was also published in the German language.

Publications on empirical studies performed in Germany were diverse and appeared to concentrate on the use of validity measures in clinical and rehabilitation contexts (e.g., Kobelt-Pönicke et al., 2020; Merten et al., 2020) as well as in forensic patients with psychiatric diagnoses (e.g., Stevens et al., 2018).

## Great Britain

The position described for Great Britain (GB) in the 2013 review paper was that the majority of research focused on PVTs, and most studies utilized non-forensic clinical populations. There has been a paucity of test validation research since this time, but again the few studies that have been conducted have used non-forensic clinical populations (Hampson et al., 2014; Suesse et al., 2015).

It would seem that clinicians in the GB continue to adopt a softer approach to SVA compared with North America. They are more reluctant to use PVTs and SVTs to identify malingering as reflected in the GB research, which has a paucity of studies using known-group designs with “malingering” groups. Clinicians in GB may be more skeptical about formulating opinions or beliefs about “malingering” due to malingering being a decision for the Courts to decide and not a clinical decision for an expert witness. There is also evidence that GB clinicians still prefer the use of the term effort test rather than PVT (e.g., Hampson et al., 2014; McGuire et al., 2019).

McWhirter et al. (2020) reviewed PVT failure in clinical populations. The authors hypothesized that PVTs measure a range of factors including attentional deficit. Larrabee et al. (2020) criticized this review, and McWhirter et al. also responded to their criticisms. In this exchange, a difference between the GB and US use of terminology and opinion about PVTs was highlighted, particularly with regards to the term effort. Larrabee stated that the term effort tests are no longer in use in the US, in part because PVTs require little effort to perform so that people experiencing significant cognitive impairment can pass them. They highlight a problem with using this term and state “continuing to refer to PVTs as “effort tests” allows mischaracterization of PVTs as sensitive attentional tasks affected by variable “effort” rather than measures of performance validity that are failed due to invalid test performance. There was some more discussion about the proper use of terminology in Britain, which can be downloaded from the journal website.

In the previous review, Merten et al. (2013) noted that neuropsychologists remained skeptical about the use of PVTs, although the majority of neuropsychologists were using them in medico-legal settings (McCarter et al., 2009). The trend continues, and they are still not widely adopted in clinical settings (Suesse et al., 2015).

Since the last review, no further detailed studies have been published which review whether neuropsychologists’ practice has changed in GB. However, there have been other reviews which have outlined the frequency in which psychologists and other professionals use SVTs/PVTs. Cartwright et al. (2019) found that only 20% of expert witness psychologists used SVTs in non-cognitive psychological assessments. Allcott et al. (2014) surveyed the practices of a range of GB expert witnesses in the fields of neurology, neuropsychiatry, neurology, orthopedics, neuropsychology, clinical psychology, and care. They found 49% of expert witnesses evaluated symptom validity by making judgments about whether there were marked inconsistencies between complaints and medical history. Thirty-two percent assessed it by determining whether complaints were disproportionate to the severity of the injury. Forty-four percent of respondents did not routinely use any tests/procedures for symptom validation. Half

of those who routinely use some form of symptom validity testing did not specify any peer-reviewed sources that were useful in their practice (i.e., 55% of respondents did not reply when asked to specify peer-reviewed articles or books that they found useful on the subject). In response to the use of these methods, the expert witnesses made comments such as the “validity of such instruments remains questionable;” “I am unaware of any reliable tests or procedures that are of help;” “I have found personal experience more useful than any of the above (peer-reviewed publications).” The review concluded that the overall impression is that most experts, including very seasoned experts, remain skeptical about the use of SVTs.

GB research shows that there is a mixed acceptance that malingering or non-credible presentations are prevalent in litigant populations. Cartwright et al. (2019) found that only 9.9% of a group of 37 participating GB psychologists who conducted medicolegal assessments believed that the claimants were malingering. Allcott et al. (2014) described a substantial variation in medicolegal and psychological experts’ prevalence estimates for exaggerated or feigned health complaints. A clear majority of the respondents found that most medico-legal cases (> 75%) were presented as genuine cases, but exact numbers were not given.

In sum, the acceptance of SVA appears to be limited. GB research predominately focuses on clinical populations, and clinicians tend to be resistant to using them to detect malingering, preferring a softer approach.

## Spain

There is a clear continuity of the panorama described for Spain in Merten et al. (2013), both in the current lines of research and in the adaptation and creation of instruments, as well as in the most relevant challenges that remain. Research is ongoing in the fields of forensics (e.g., Fariña et al., 2014), neuropsychology (e.g., Daugherty et al., 2020), medicolegal (e.g., Capilla Ramírez et al., 2014), and military (e.g., García Silgo, 2019). The most prevalent field both in terms of research and application is forensic assessment, in particular the assessment of sequelae of psychological injuries subsequent to traumatic events (like gender-based violence, Marín-Torices et al., 2018, or of traffic accidents, Puente-López et al., 2021).

Spanish adaptations of a variety of international validity tests are available (MMPI-2, MMPI-A, MMPI-2-RF, Personality Assessment Inventory [PAI, PAI-A], SIMS, Test of Memory Malingering [TOMM]), and continue to be used in research studies in Spain (e.g., López-Miquel & Pujol-Robinat, 2020; Vilar-López et al., 2021). The Spanish adaptation of the MMPI-A-RF has recently been published, and the MMPI-3 publication is planned for 2023. Specific malingering scales for forensic assessment of posttraumatic

stress disorder were developed in Spain, such as the Trauma Impact Questionnaire (CIT; Crespo et al., 2020) and the Posttraumatic Stress Disorder Symptom Severity Scale: Forensic version (EGS-F; Echeburúa et al., 2017). Performance validity tests have also been created in the field of neuropsychology, like the extended version of the Coin-in-the-Hand Test, developed in Spain and later validated at a multicultural level (Daugherty et al., 2021).

Despite the availability of such research and instruments, further basic and applied research is necessary. To date, there appear to be more open questions than answers. Similar to the scenario depicted in 2013, it is still necessary to define adequate protocols for the systematic investigation of possible malingering based on consensus across the different fields of application and areas of assessment. Unfortunately, this goal appears to be quite distant. Furthermore, similar to the situation depicted above for Germany, the complexity of validity assessment is still underestimated and downplayed; the search for a “magic wand” of simple and fast solutions persists. In this sense, there continues to be an inadequate use of screening tools (like the SIMS); such instruments are partly used for diagnostic purposes with a poor understanding of their scope and limitations. The medicine field is facing a special challenge with regard to the assessment of temporary disability due to mental health disorders. This area certainly requires further research and elaboration, more profound professional specialization, and the improvement of assessment protocols.

## The Netherlands

In the Netherlands, SVA has attracted steady research attention although the field was hardly supported by National or European grant organizations. Since 2013, 6 doctoral theses on symptom validity have been published (Boskovic, 2019; Dandachi-FitzGerald, 2017; Meyer, 2020; Niesten, 2019; Van der Heide, 2021; Van Impelen, 2018). [All doctoral theses are accessible, see reference list]. In addition to deception detection, Dutch research on SVA is characterized by conceptual studies (e.g., Merckelbach et al., 2019). Experimental studies have examined whether moral primes (Niesten et al., 2017) and feedback (Merckelbach et al., 2015) can deter symptom overreporting tendencies. Also, studies have looked into the consequences of symptom and performance invalidity (e.g., Merckelbach et al., 2014a; Roor et al., 2021).

Two new performance validity measures have been developed; the Groningen Effort Test, an attention-based performance validity test (Fuermaier et al., 2017), and the Visual Association Test–Extended, a memory test with an embedded performance validity index (Meyer et al., 2017). Additionally, Dutch versions of the Assessment of Depression Inventory (ADI-NL; Mogge & LePage, 2004; Van Leeuwen

& de Jonghe, 2018) and the Schretlen Malingering Scale (Merckelbach, Otgaar et al., 2014b; Schretlen et al., 1992) have been made available.

In the professional field, the issue of validity assessment has raised increasing interest among insurance and company doctors, as well as among lawyers, especially those specialized in personal injury claims. Like in other countries, the new nomenclature of distinguishing performance and symptom validity tests (measuring underperformance on cognitive tests and overreporting of symptoms, respectively) has been adopted. Terms like “malingering test” and “effort test” are less commonly used, and effort in the context of performance validity is more clearly understood as “applying effort to perform well.” The revised guideline for forensic neuropsychological assessments now explicitly states that “in every forensic neuropsychological assessment, the evaluation of symptom and performance validity must be psychometrically substantiated” (Nederlands Instituut voor Psychologen, sectie Neuropsychologie, 2016, p.10, quotation translated). This guideline further stipulates that a minimum of two freestanding validity tests should be administered, and that performance and symptom validity should be separately assessed.

In contrast, the idea that clinical impression suffices to assess the validity of self-reported symptoms is still commonly voiced among forensic psychiatrists. According to their guideline, psychiatrists may consider the use of specific instruments as soon as they have doubts about symptom validity based on their clinical impression (Nederlandse Vereniging voor Psychiatrie, 2012). This primacy of clinical judgment flies in the face of what is now becoming an impressive corpus of knowledge (e.g., Dandachi-FitzGerald et al., 2017; Rosen & Phillips, 2004; Zuberá et al., 2015). Up until now, studies on how frequently validity tests are used in forensic assessments in the Netherlands are lacking.

To examine the role of symptom validity tests in a Dutch court, Merckelbach and Dandachi-FitzGerald (2021) searched the public database on court decisions with the terms “feigning,” “simulation,” “malingering,” and “exaggeration,” and selected the ten most recent decisions for each term. In 22 of the 36 cases (61%), validity tests were mentioned; showcasing that by now these tests have acquired a fixed position in the legal system. Still, a close analysis of these legal cases revealed that there is considerable room for improvement, specifically when it comes to interpreting the outcomes of validity tests. For example, in one case, poor performance validity was explained by the psychiatrist as “unconscious exaggeration caused by a conversion disorder.” In yet another case, the neuropsychologist concluded that the failures on two freestanding performance validity tests could be explained away by cognitive deficits due to mild TBI. In the first case, the dubious explanation was accepted by the court. In the second case, the court rightly ruled that

the expert opinion on validity test failure was incoherent because symptom validity test failure casts doubts on the possibility to establish the presence of cognitive deficits with any acceptable degree of certainty. These cases illustrate the importance of both experts and judges being well informed about SVA; there is still much work to do here.

To conclude, the state-of-the-art of SVA in the Netherlands appears to be at the forefront of Europe, at least as far as it concerns neuropsychological assessments. Nonetheless, controversies remain and pertain mostly to the interpretation of validity test outcomes. Experts struggle with how to interpret a patient’s symptom presentation when this patient passes some validity tests but fails others. Also, there still seems to be an inclination to ascribe validity test failure to psychopathology or somatic symptoms such as fatigue and pain, highlighting that problematic beliefs about SVA are circling around.

## Austria

The Austrian Federal Ministry of Health has published guidelines for the preparation of clinical-psychological and health-psychological data and reports. According to the Psychologists ‘ Act, these guidelines are binding for psychologists (Bundesministerium für Soziales, Gesundheit, Pflege, und Konsumentenschutz 2020). The problem of symptom and performance validity is not specifically addressed in the current guidelines. However, this issue has repeatedly been taken up to varying degrees of detail in recent publications by Austrian authors (Lehrner et al., 2015, 2021; Lettner, 2019; Strubreither, 2021).

The range of advanced training courses addressing SVA has improved significantly in Austria over the past few years. Workshops on the subject of “symptom and performance validation in neuropsychological reports” have been organized and are still being offered by the Austrian Neuropsychological Association (GNPÖ). The content of these workshops covers the interpretation and misinterpretation of results, the possibilities and limits of the use of a specific test to detect malingering, ethical questions raised by the use of SVTs and PVTs, and the presentation and discussion of expert reports. Also, the curriculum for legal psychology of the Professional Association of Austrian Psychologists (BÖP) offers a training module (Module 2) comprising special topics such as symptom exaggeration, malingering, and symptom validation. In May 2021, an online advanced training course on SVA in clinical psychological assessment was held via Zoom as part of the advanced training of the clinical psychology section of the BÖP. It attracted approximately 500 participants.

A questionnaire on SVA in psychological assessment was sent out by the BÖP to all participants. It was also sent to all 5054 members of the clinical psychology section of

the BÖP, as well as 87 expert psychologists listed as court experts of the Federal Ministry of Justice. A total of 99 submitted data sets could be analyzed. These data sets stemmed from 17 psychologists listed as experts at the regional courts and 82 members of the clinical psychology section. The two groups differed in the reported frequency of validity test use. Sixteen percent of the section members and 12% of the court experts reported that they used validity tests in more than 95% of their clinical assessment cases. In an independent psychological examination, 25% of the section members and 29% of the experts stated that they used SVTs and PVTs in more than 95% of their cases. A 0% of court experts reported that they never used validity tests. In contrast, 28% of section members reported that they never used PVTs/SVTs for clinical cases, and 13% reported that they never used PVTs/SVTs in court-ordered examinations. The frequency of PVT/SVT use is similar to that which is reported in other European countries (e.g., Dandachi-FitzGerald et al., 2013), but section members were not using PVTs/SVTs as frequently as psychologists in forensic contexts did.

It is noteworthy that there is a prominent European publisher and distributor of computerized psychological assessment, Schuhfried, which is based in Austria. Among others, they published the Groningen Effort Test (Fuermaier et al., 2017).

Among research activities in the field of SVA, a clinical study by Bodner et al. (2019) investigated the validity of several PVTs (TOMM, Fifteen-Item Test, Reliable Digit Span, and Reliable Spatial Span) in the context of language disorders (aphasia). At the Medical University of Vienna, Czornik et al. (2021) evaluated a range of tests (e.g., Word Memory Test, the SIMS, and SRSI) using a sample of individuals from a memory out-patient clinic. A further study by Czornik et al. (2022, in this issue) investigated a reaction-time-based embedded PVT in a sample of civil forensic patients.

Symptom and performance validation continues to be discussed controversially in Austria. With no strict relevant guidelines available, individual professionals approach this topic differently. Yet, with more widespread knowledge about SVA, the use of PVTs and SVTs both in clinical and in forensic contexts is increasing.

## Italy

A survey of SVA practices and beliefs of Italian psychologists was conducted recently by Giromini et al. (under review). According to that survey, the majority of Italian practitioners (> 60%) are prone to use SVTs and/or PVTs when they believe that their evaluatee could have an interest in producing false or grossly exaggerated physical or psychological symptoms. However, only 13.2% reported using one or more stand-alone SVTs or PVTs routinely in their

assessments. Accordingly, Giromini and colleagues concluded that, albeit Italian psychologists do not always question the credibility of presented symptoms, when they do so, they are relatively prone to use SVTs and/or PVTs to assist their decision-making.

With regard to research, a simple literature search found 48 articles potentially focused on SVA in Italy.<sup>1</sup> Of these 48, eleven were not directly relevant to our project, as they focused either on underreporting (e.g., Pompili et al., 2003; Roma et al., 2018) or on other loosely related issues (e.g., on the difficulties in completing a literature review on attention deficit hyperactivity disorder in adulthood, given the influence of multiple factors, including malingering; Mucci et al., 2018). Of the remaining 37 articles, as many as 26 (70%) were published during the past 5 years alone (i.e., between 2016 and 2021), thus highlighting an ongoing growing interest in SVA within the Italian context.

These most recent research efforts primarily focused on three major topics. First, a few Italian researchers investigated the potential usefulness of various, modern technological advancements. For instance, Orrù et al. (2021) and Pace et al. (2019) applied machine-learning techniques to develop a shorter version of the SIMS (Orrù) and to discriminate credible from noncredible presentations using the b test (Pace). Monaro et al. (2018) analyzed mouse movements of individuals either feigning depression or responding honestly while engaged in a double-choice computerized task, so as to develop a machine learning-based algorithm aimed at detecting feigned depression. Zago et al. (2019) implemented facial thermography and kinematic analyses, in addition to symptom validity testing, in an effort to help detection of feigned amnesia after committing a crime.

A second emerging research area in Italy concerns the investigation of the effectiveness of several SVTs and PVTs. In particular, numerous recent studies examined the psychometric properties of the Italian IOP-29, reporting on its concurrent (Giromini et al., 2018), incremental (Giromini et al., 2019), and ecological (Roma et al., 2020) validity, on its applicability to multiple symptom presentations (Giromini et al., 2020), and on the equivalence of its online and paper-and-pencil formats (Giromini et al., 2021). Additionally, some authors also investigated the effectiveness of the SIMS and MMPI-2 in detecting noncredible presentations (e.g., Mazza et al., 2019), and a recently published article described the development and initial validation of the IOP-M, a new, add-on, PVT module designed to be used

<sup>1</sup> This literature search was conducted through PsychInfo on May 11, 2021, using the terms “symptom validity,” “performance validity,” “malingering,” or “feigning,” together with the terms “Italy” or “Italian.”

in combination with the IOP-29 (see also Banovic et al., 2021; Carvalho et al., 2021; Gegner et al., 2021).

Lastly, a third research area that deserves mention here concerns the detection of feigned crime-related amnesia. An Italian study investigated whether feigning amnesia for a mock crime has an impact on an individual's ability to later recall the actual details of the mock crime (Mangiulli et al., 2018).

It should be noted, however, that Italian research on SVA actually goes beyond these three research areas. For instance, some relatively recent Italian publications addressed the feigning of specific problems such as second-language deficit subsequent to mild traumatic brain injury (Zago et al., 2013) or elaborated on malingering-related conditions such as the factitious disorder (Poloni et al., 2019) or Munchausen syndrome (Callegari et al., 2006). In fact, as noted above, Italian research on SVA-related topics is accumulating rapidly, and one may anticipate that this trend will likely continue during the coming years.

## Switzerland

The situation of SVA in Switzerland in recent years was characterized by an increasing acceptance of the fact that it is a useful and necessary tool to distinguish valid symptoms from invalid (exaggerated or feigned) complaints. In 2008, the Swiss Federal Social Insurance Office commissioned and published a study (Kool et al., 2008) with the aim of providing a systematic review of the literature on SVA to promote the development and adoption of medico-legal standards among professionals. Also, in the guidelines for medico-legal neuropsychological assessment of the Swiss Association of Neuropsychologists (SVNP, 2011), testing of effort and determinations about the consistency of test results were described as integral parts of a neuropsychological examination in a medico-legal context. This was confirmed to be necessary for the legal literature (Kieser, 2012); accordingly, Swiss courts increasingly emphasized the importance of SVA in relevant judicial decisions. Plohmann and Hurter (2017) published the first study to examine the prevalence of inadequate effort and malingered neurocognitive dysfunctions in medico-legal contexts in Switzerland. The authors reported a prevalence of probable or definite malingered neurocognitive dysfunction in medico-legal contexts ranging from 27.5 to 34.3%, depending upon which cut score was used for Reliable Digit Span. Within this group, about one-tenth (10.3–12.8%) presented with below-chance response patterns and qualified as cases of definite malingering. The prevalence rates in Switzerland were in line with those obtained in other countries (e.g., Mittenberg et al., 2002) and demonstrated the necessity of performing a careful SVA in medico-legal evaluations. The fifth European Conference on Symptom Validity Assessment was held in Basel in 2017. With special

emphasis on psychosomatic, psychiatric, and pain disorders, it was organized under the auspices of the Swiss Association of Neuropsychologists.

Current efforts are directed at training experienced and young neuropsychologists in the use and interpretation of neuropsychological tests, taking into account adequate SVA. Postgraduate training for what is called “Eidgenössisch anerkannter Neuropsychologietitel (EAN)” and for a “Master of Advanced Studies in Neuropsychology (MAS)” was established at the University of Zurich in 2020. A specific module is dedicated to SVA. Moreover, SVA is also one central topic in several modules of the Swiss Insurance Medicine (SIM) assessor training leading to the qualification as a “certified neuropsychological assessor SIM.” A “SIM specialist group in neuropsychology” was founded in 2020. In the coming years, further efforts are needed to establish high qualitative standards in SVA both in medico-legal and in clinical fields.

## Major Challenges for Future Developments

It may be stated that research into and forensic practice of symptom and performance validity testing in Europe, seen as a whole, has developed at a level comparable to that known from the U.S. and Canada. Yet, a closer look reveals a continued gross heterogeneity across the continent. Another comprehensive survey on SVT/PVT use, following the Dandachi-FitzGerald et al. (2013) study, with the inclusion of as many national neuropsychological societies as possible appears to be indicated for the years to come, in order to tap the state of the art that will be arrived at in the course of the 2020s. In some countries, forensic and partly clinical practice underwent a significant change with the more widespread use of validity measures, but there is little or no information about other parts of the continent.

In comparison to the situation about 10 years ago, not only a significant body of empirical studies has been accumulated, but also conceptual and practical aspects of SVA underwent significant modifications. Challenges arise for practitioners to always keep abreast of methodological and conceptual developments at the highest level of current knowledge. As described in some of the national reports above, the conceptual shift from “malingering research” and “malingering detection” to “validity research” and “validity assessment” is not readily embraced by all researchers and all practitioners, and outstanding position papers like Sherman et al. (2020) and Sweet et al. (2021) will not be absorbed quickly and smoothly in all corners of the continent. Between overt neurological disease/brain damage and frank malingering, there are many other conditions, including the exaggeration of minor neurological injury and psychiatric conditions such as factitious disorder, somatoform conditions, and what is now called functional neurological

symptom disorder (cf. Stone & Sharpe, 2020, for a recent appraisal of the latter). Thus, related to conceptual developments, it is necessary to further explore questions regarding what validity failures actually mean in different contexts and what the legal and treatment implications are. In interdisciplinary settings, this may require exploration by a range of disciplines. This is particularly relevant for psychiatric conditions such as post-traumatic stress, somatoform disorders, and pain-related disabilities (e.g., Greve et al., 2012; Howe, 2012; Merten & Merckelbach, 2013). In some conditions where there is an overlap between diagnostic categories (e.g., somatoform and conversion disorders, factitious disorder, and malingering), it can be a problem pigeonholing patients into one of them as individuals may equally fit into more than one diagnostic category (e.g., Merten & Merckelbach, 2013; Sherman et al., 2020). Different conditions may co-occur, with no clear boundaries, but smooth transitions between them.

On the methodological level, validity research will have to move further away from the easy-to-do analog studies into real-world settings, in particular with well-defined clinical patient groups. However, the primary challenge in such settings is that it will be difficult, if not impossible, to reliably tell apart true-positive from false-positive SVT or PVT results in some constellations (e.g., Dandachi-FitzGerald et al., 2016; Merten et al., 2020). On the level of test development, professionals in only a few European countries with non-English national languages appear to dispose upon a sufficient number of well-validated SVTs and PVTs, most of them adaptations of North American tests. For some nations, the availability of tests is a major problem (e.g., Janaviciute et al., 2021). Also, equivalence studies comparing different language versions are rare. A focus on European tests (e.g., Meyer et al., 2017; Walter et al., 2016) will certainly not solve the basic problems posed by the diverse range of languages and cultures that are present across the continent. The continuing influx of immigrants from Asia and Africa is another factor aggravating intercultural problems of validity assessment. There is a clear need for multi-language versions of common validity measures. On the level of test administration, modernization and recent restrictions due to the COVID-19 pandemic have fostered online presentation modes of tests (remote assessment), with yet unknown consequences for the interpretation/interpretability of SVTs and PVTs outside their standard conditions of use. The validity of these tests has not yet been systematically researched outside of the normal use. Also, before psychologists can use a test remotely, the copyright holder of the instrument must agree to their test being used in this manner. To our knowledge, only one European study has addressed the question of paper–pencil versus online presentation to date (Giromini et al., 2021). It is, therefore, necessary to conduct more systematic research into these problems.

Another special challenge to continue in the future is to further educate practitioners to correctly use and interpret the results of validity testing, in particular, to resist temptations to explain away uncomfortable results of validity assessment (e.g., Dandachi-FitzGerald et al., 2015; Merten, 2017). In most countries, proper in-depth routine training in methods of SVA is often omitted both for neuropsychologists and for forensic psychologists.

Professional guidelines for forensic assessment and for independent medical and psychological evaluations appear to include increasingly statements about symptom and performance validation, but special guidelines are rare. Those published in Britain (McMillan et al., 2009) have recently been updated (Moore et al., 2021). Another important issue is research and guidelines on how to handle clinical patients who produce invalid test profiles or report noncredible symptoms (e.g., Carone & Bush, 2018; Martin & Schroeder, 2021).

A larger number of open questions and important problems can easily be identified; consequently, with the continued relevance of the topic, research activities are likely not to slow down in the foreseeable future. The study of both professionals' and laypersons' attitudes and expectations with regard to factitious symptom presentations, malingering, fraudulent health claims, etc. will be another problem of interest, not least with respect to social and intercultural factors (e.g., Cartwright & Roach, 2015; Dandachi-FitzGerald et al., 2020; Merten & Giger, 2018; Schlicht & Merten, 2014). Also, embedded PVTs are clearly underresearched in Europe, contrary to their apparent significance in validity research and practice. Similarly, the use of multiple validity measures and their consequences for diagnostic decision-making is underresearched. In contrast to research activities in other parts of the world, validity assessment with personality inventories, in particular the MMPI family and the Personality Assessment Inventory, appear to play a minor role in Europe (with some exceptions, e.g., García Silgo, 2019; Giromini et al., 2019; Vossler-Thies et al., 2013). Remote assessment and its consequences for validity is certainly another topic of interest, in particular, if the COVID-19 crisis continues to affect professional activities as much as it did in 2020 and 2021 (Corey & Ben-Porath, 2020). In another 10 years' time, we will certainly know more about these topics, and others will have emerged not even mentioned in this review.

## References

- Ægisdóttir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., Cook, R. S., Nichols, C. N., Lampropoulos, G. K., Walker, B. S., Cohen, G., & Rush, J. D. (2006). The meta-analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction.



- The Counseling Psychologist*, 34(3), 341–382. <https://doi.org/10.1177/0011000005285875>
- Allcott, D., Anderson, S., Friedland, D., Leng, N., Gross, M., Skelton-Robinson, M., & Weller, M. (2014). How do experts reporting for the legal process validate symptoms? The results of a survey. *Medicine, Science and the Law*, 54(2), 68–73. <https://doi.org/10.1177/0025802413491247>
- Anderson, S. (2010). Conference report: The first european symposium on symptom validity assessment (May 8–9 2009). *Zeitschrift Für Neuropsychologie*, 21(1), 65. <https://doi.org/10.1024/1016-264X/a000007>
- Banovic, I., Filippi, F., Viglione, D. J., Scrima, F., Zennaro, A., Zappalà, A., & Giromini, L. (2021). Detecting coached feigning of schizophrenia with the inventory of problems–29 (IOP-29) and its memory module (IOP-M): A simulation study on a French community sample. *International Journal of Forensic Mental Health*. <https://doi.org/10.1080/14999013.2021.1906798>
- Bodner, T., Merten, T., & Benke, T. (2019). Performance validity measures in clinical patients with aphasia. *Journal of Clinical and Experimental Neuropsychology*, 41(5), 476–483. <https://doi.org/10.1080/13803395.2019.1579783>
- Boskovic, I. (2019). *A multi-method approach to the detection of fabricated symptoms*. Doctoral thesis. Maastricht University. <https://doi.org/10.26481/dis.20190129ib>
- Bundesministerium für Soziales, Gesundheit, Pflege und Konsumentenschutz. (2020). Gutachterrichtlinie – Richtlinie für die Erstellung von klinisch-psychologischen und gesundheitspsychologischen Befunden und Gutachten [Guidelines for clinical-psychological reports and expert reports]. Download from [https://www.sozialministerium.at/dam/jcr:c1366581-f962-42b4-abb6-a7ef39a40f30/Gutachterrichtlinie%20\(Stand%202020%20Juni\).pdf](https://www.sozialministerium.at/dam/jcr:c1366581-f962-42b4-abb6-a7ef39a40f30/Gutachterrichtlinie%20(Stand%202020%20Juni).pdf)
- Callegari, C., Bortolaso, P., & Vender, S. (2006). A single case report of recurrent surgery for chronic back pain and its implications concerning a diagnosis of Münchhausen syndrome. *Functional Neurology*, 21(2), 103–108.
- Capilla Ramírez, P., González Ordi, H., Santamaría, P., & Casado Morales, M. I. (2014). Detección de exageración de síntomas en esguince cervical: Pacientes clínicos versus sujetos análogos [Detection of symptom exaggeration in whiplash: Patients versus analogue participants]. *Trauma*, 25(1), 4–10.
- Carone, D. A., & Bush, S. S. (2018). *Validity assessment in rehabilitation psychology and settings*. Oxford University Press.
- Cartwright, A., & Roach, J. (2015). Fraudulently claiming following a road traffic accident: A pilot study of UK residents' attitudes. *Psychiatry, Psychology and Law*, 23(3), 446–461. <https://doi.org/10.1080/13218719.2015.1080148>
- Cartwright, A., Roach, J., & Armitage, R. (2019). Mission impossible? Assessing the veracity of a mental health problem as result of a road traffic accident: A preliminary review of UK experts' practices. *Journal of Forensic Psychiatry & Psychology*, 30(1), 89–111. <https://doi.org/10.1080/14789949.2018.1502338>
- Carvalho, L., Reis, A., Colombarolli, M. S., Pasian, S. R., Miguel, F. K., Erdodi, L. A., Viglione, D. J., & Giromini, L. (2021). Discriminating feigned from credible PTSD symptoms: A validation of a Brazilian version of the inventory of problems–29 (IOP-29). *Psychological Injury and Law*, 14(1), 58–70. <https://doi.org/10.1007/s12207-021-09403-3>
- Corey, D. M., & Ben-Porath, Y. S. (2020). Practical guidance on the use of the MMPI instruments in remote psychological testing. *Professional Psychology: Research and Practice*, 51(3), 199–204. <https://doi.org/10.1037/pro0000329>
- Crespo, M., González-Ordi, H., Gómez-Gutiérrez, M., & Santamaría, P. (2020). *CIT, Cuestionario de Impacto del Trauma* [Trauma Impact Questionnaire]. TEA Ediciones.
- Czornik, M., Merten, T., & Lehrner, J. (2021). Symptom and performance validation in patients with subjective cognitive decline and mild cognitive impairment. *Applied: Neuropsychology Adult*, 28(3), 269–281. <https://doi.org/10.1080/23279095.2019.1628761>
- Czornik, M., Seidl D., Tavakoli, S., Merten, T., & Lehrner, J. (2022, in this issue). Motor reaction times as an embedded measure of performance validity: A study with a sample of Austrian early retirement claimants. *Psychological Injury and Law*.
- Dandachi-FitzGerald, B. (2017). *Symptom validity in clinical assessments*. Doctoral thesis. Maastricht University. <https://doi.org/10.26481/dis.20170602bdf>
- Dandachi-FitzGerald, B., Merckelbach, H., Bošković, I., & Jelicic, M. (2020). Do you know people who feign? Proxy respondents about feigned symptoms. *Psychological Injury and Law*, 13(3), 225–234. <https://doi.org/10.1007/s12207-020-09387-6>
- Dandachi-FitzGerald, B., Merckelbach, H., & Ponds, R. W. H. M. (2015, June). *Neuropsychologists' ability to predict distorted symptom presentation and professional communication of SVT failure*. Paper presented at the Fourth European Conference on Symptom Validity Assessment. Maastricht, The Netherlands.
- Dandachi-FitzGerald, B., Merckelbach, H., & Ponds, R. W. H. M. (2017). Neuropsychologists' ability to predict distorted symptom presentation. *Journal of Clinical and Experimental Neuropsychology*, 39(3), 257–264. <https://doi.org/10.1080/13803395.2016.1223278>
- Dandachi-FitzGerald, B., Ponds, R. W. H. M., & Merten, T. (2013). Symptom validity and neuropsychological assessment: A survey of practices and beliefs of neuropsychologists in six European countries. *Archives of Clinical Neuropsychology*, 28(8), 771–783. <https://doi.org/10.1093/arclin/act073>
- Dandachi-FitzGerald, B., van Twillert, B., van de Sande, P., van Os, Y., & Ponds, R. W. H. M. (2016). Poor symptom and performance validity in regularly referred hospital outpatients: Link with standard clinical measures, and role of incentives. *Psychiatry Research*, 239, 47–53. <https://doi.org/10.1016/j.psychres.2016.02.061>
- Daugherty, J. C., Puente, A. E., Lozano-Ruiz, A., & Perez-Garcia, M. (2020). Cultural accommodations for cutoff scores of embedded performance validity tests in a Spanish college population. *Applied Neuropsychology: Adult*. <https://doi.org/10.1080/23279095.2020.1846536>
- Daugherty, J. C., Querido, L., Quiroz, N., Wang, D., Hidalgo-Ruzzante, N., Fernandes, S., Pérez-García, M., De Los Reyes-Aragon, C. J., Pires, R., & Valera, E. (2021). The coin in hand—extended version: Development and validation of a multicultural performance validity test. *Assessment*, 28(1), 186–198. <https://doi.org/10.1177/1073191119864652>
- Deutsche Rentenversicherung. (2017, 11. Oktober). Deutsche Rentenversicherung begrüßt erste Urteile gegen „Rentenbetrüger“. [http://www.deutsche-rentenversicherung.de/Westfalen/de/Inhalt/4\\_Presse/Journalisten/Pressemitteilungen/2017/Urteil\\_gegen\\_Rentenbetrueger\\_11\\_10\\_2017.html](http://www.deutsche-rentenversicherung.de/Westfalen/de/Inhalt/4_Presse/Journalisten/Pressemitteilungen/2017/Urteil_gegen_Rentenbetrueger_11_10_2017.html)
- Dressing, H., Foerster, K., Widder, B., Schneider, F., & Falkai, P. (2011). Zur Anwendung von Beschwerdenuvalidierungstests in der psychiatrischen Begutachtung [On the application of symptom validity tests in psychiatric forensic examination]. *Der Nervenarzt*, 82(3), 388–390. <https://doi.org/10.1007/s00115-011-3275-x>
- Echeburúa, E., Amor, P. J., Muñoz, J. M., Sarasua, B., & Zubizarreta, I. (2017). Escala de Gravedad de Síntomas del Trastorno de Estrés Posttraumático según el DSM-5: Versión forense (EGS-F) [Posttraumatic stress disorder symptom severity scale according to DSM-5 criteria: Forensic version (EGS-F)]. *Anuario De Psicología Jurídica*, 27(1), 67–77. <https://doi.org/10.1016/j.apj.2017.02.005>
- Fariña, F., Arce, R., Vilariño, M., & Novo, M. (2014). Assessment of the standard forensic procedure for the evaluation of

- psychological injury in intimate-partner violence. *Spanish Journal of Psychology*, 17(e32), 1–10. <https://doi.org/10.1017/sjp.2014.30>
- Fargo, J. D., Schefft, B. K., Szaflarski, J. P., Howe, S. R., Yeh, H. S., & Privitera, M. D. (2008). Accuracy of clinical neuropsychological versus statistical prediction in the classification of seizure types. *The Clinical Neuropsychologist*, 22(2), 181–194. <https://doi.org/10.1080/13854040701220093>
- Fuermaier, A., Tucha, O., Stenger, H., Koerts, J., Aschenbrenner, S., & Tucha, L. (2017). *The Groningen effort test (GET)*. Test manual. Schuhfried.
- García Silgo, M. (2019). Detección de simulación de trastorno mental mediante el MMPI-2-RF, el PAI y el SIMS: Estudio de análogos en una muestra militar [Mental disorder malingering detection with MMPI-2-RF, PAI, and SIMS in a military analog design]. *Sanidad Militar*, 75(4), 203–213. <https://doi.org/10.4321/s1887-85712019000400004>
- Gegner, J., Erdodi, L. A., Giromini, L., Viglione, D. J., Bosi, J., & Brusadelli, E. (2021). An Australian study on feigned mTBI using the inventory of problems–29 (IOP-29), its memory module (IOP-M), and the Rey fifteen item test (FIT). *Applied Neuropsychology: Adult*. <https://doi.org/10.1080/23279095.2020.1864375>
- Giromini, L., Carfora Lettieri, S., Zizolfi, S., Zizolfi, D., Viglione, D. J., Brusadelli, E., Perfetti, B., di Carlo, D. A., & Zennaro, A. (2019). Beyond rare-symptoms endorsement: A clinical comparison simulation study using the Minnesota multiphasic personality inventory-2 (MMPI-2) with the inventory of problems–29 (IOP-29). *Psychological Injury and Law*, 12(3–4), 212–224. <https://doi.org/10.1007/s207-019-09357-7>
- Giromini, L., Pignolo, C., Young, G., Drogin, E. Y., Zennaro, A., & Viglione, D. J. (2021). Comparability and validity of the online and in-person administrations of the inventory of problems-29. *Psychological Injury and Law*, 14(1), 77–88. <https://doi.org/10.1007/s12207-021-09406-0>
- Giromini, L., Viglione, D. J., Pignolo, C., & Zennaro, A. (2018). A clinical comparison, simulation study testing the validity of SIMS and IOP-29 with an Italian sample. *Psychological Injury and Law*, 11(4), 340–350. <https://doi.org/10.1007/s12207-018-9314-1>
- Giromini, L., Viglione, D. J., Pignolo, C., & Zennaro, A. (2020). An Inventory of Problems–29 Sensitivity study investigating feigning of four different symptom presentations via malingering experimental paradigm. *Journal of Personality Assessment*, 102(4), 563–572. <https://doi.org/10.1080/00223891.2019.1566914>
- 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). Springer Publ.
- Greve, K. W., Bianchini, K. J., & Brewer, S. T. (2012). The assessment of performance and self-report validity in persons claiming pain-related disability. *The Clinical Neuropsychologist*, 27(1), 108–137. <https://doi.org/10.1080/13854046.2012.739646>
- Hampson, N. E., Kemp, S., Coughlan, A. K., Moulin, C. J., & Bhakta, B. B. (2014). Effort test performance in clinical acute brain injury, community brain injury, and epilepsy populations. *Applied Neuropsychology: Adult*, 21(3), 183–194. <https://doi.org/10.1080/09084282.2013.787425>
- Hoffmann, A. (2019). Mit Betrug in der Pflege verdient man wie im Drogenhandel [Fraud in disability care is financially as lucrative as drug dealing]. <https://www.stern.de/wirtschaft/news/mit-betrug-in-der-pflege-verdient-man-wie-im-drogenhandel-8570136.html>, Download Nov. 17, 2020.
- Howe, L. L. S. (2012). Distinguishing genuine from malingered posttraumatic stress disorder in head injury litigation. In C. R. Reynolds & A. M. Horton (Eds.), *Detection of malingering during head injury litigation* (2nd ed., pp. 301–331). Springer.
- Janaviciute, J., Lelyte, I., & Zukevicius, A. (2021). Rey penkiolikos objektu testo verte atpazistant simuliavusiusius atminties sutrikismus [Utility of the Rey 15-Item Test for detecting memory malingering]. *Psichologija*, 63, 8–23. <https://doi.org/10.15388/Psichol.2021.23>
- Kahneman, D., Sibony, O., & Sunstein, C. R. (2021). *Noise: A flaw in human judgment*. Little, Brown Spark.
- Kieser, U. (2012). Neuropsychologie – Stellenwert und Bedeutung in der sozialversicherungsrechtlichen Rechtsprechung des Bundesgerichts [Neuropsychology: Its role and significance in social-law jurisdiction of the Swiss federal court]. In U. Kieser & M. Lendfers (Eds.), *Jahrbuch zum Sozialversicherungsrecht* (pp. 167–178). Dike Verlag.
- Kobelt-Pönicke, A., Walter, F., & Riemann, M. (2020). Führt das Bewusstsein moralischer Grundwerte zu einem authentischeren Antwortverhalten in Beschwerdenvalidierungstests? [Does emphasizing moral values decrease dishonest answers in symptom validity tests?] *Zeitschrift für Psychiatrie, Psychologie Und Psychotherapie*, 68(2), 106–112. <https://doi.org/10.1024/1661-4747/a000409>
- Kool, J., Meichtry, A., Schaffert, R., & Rüesch, P. (2008). *Der Einsatz von Beschwerdevalidierungstests in der IV-Abklärung: Bericht im Rahmen des mehrjährigen Forschungsprogramms zu Invalidität und Behinderung (FoP-IV)* [Use of symptom validity tests in disability examinations]. Bundesamt für Sozialversicherungen.
- Larrabee, G. J. (2012). Performance validity and symptom validity in neuropsychological assessment. *Journal of the International Neuropsychological Society*, 18(4), 625–631. <https://doi.org/10.1017/S1355617712000240>
- Larrabee, G. J., Boone, K. B., Bianchini, K. J., Rohling, M. L., & Sherman, E. M. (2020). *Response to McWhirter et al.* Discussion material, linked to the McWhirter et al. (2020) article at publisher's website, <https://jnnp.bmj.com/content/91/9/945>
- Lehrner, J., Bodner, T., Krafack, A., & Merten, T. (2015). Psychologische Begutachtung in Österreich [Independent psychological assessment in Austria]. *Psychologie in Österreich*, 5(5), 368–376.
- Lehrner, J., Neuwirth, W., Fanninger, R., Krafack, A., & Sattler, W. (2021). Klinisch-(neuro)psychologische Leistungsbeurteilung im Gutachten [Psychological and neuropsychological assessment of performance in expert examinations]. *Österreichische Zeitschrift Für Das Ärztliche Gutachten*, 9(1), 7–12.
- Lettrier, S. (2019). Das klinisch-neuropsychologische Gutachten als wissenschaftlich fundiertes Beweismittel [The neuropsychological expert report in court]. *Sachverständige*, 4, 1–3.
- López-Miquel, J., & Pujol-Robinat, A. (2020). Análisis descriptivo de la simulación de síntomas psicológicos en una muestra forense [Descriptive analysis of malingered psychological symptoms in a forensic sample]. *Revista Española De Medicina Legal*, 46(4), 175–182. <https://doi.org/10.1016/j.reml.2020.01.001>
- Mangiulli, I., Lanciano, T., Jelicic, M., van Oorsouw, K., Battista, F., & Curci, A. (2018). Can implicit measures detect source information in crime-related amnesia? *Memory*, 26(8), 1019–1029. <https://doi.org/10.1080/09658211.2018.1441421>
- Marín-Torices, M. I., Hidalgo-Ruzzante, N., Daugherty, J. C., Jiménez-González, P., & Perez Garcia, M. (2018). Validation of neuropsychological consequences in victims of intimate partner violence in a Spanish population using specific effort tests. *The Journal of Forensic Psychiatry & Psychology*, 29(1), 86–98. <https://doi.org/10.1080/14789949.2017.1339106>
- Martin, P. K., & Schroeder, R. W. (2021). Feedback with patients who produce invalid testing: Professional values and reported practices. *The Clinical Neuropsychologist*, 35(6), 1134–1153. <https://doi.org/10.1080/13854046.2020.1722243>
- Mazza, C., Orrù, G., Burla, F., Monaro, M., Ferracuti, S., Colasanti, M., & Roma, P. (2019). Indicators to distinguish symptom accentuators from symptom producers in individuals with a diagnosed adjustment disorder: A pilot study on inconsistency subtypes using SIMS and MMPI-2-RF. *PLoS One*, 14(12). <https://doi.org/10.1371/journal.pone.0227113>

- McCarter, R. J., Walton, N. H., Brooks, D. N., & Powell, G. E. (2009). Effort testing in contemporary UK neuropsychological practice. *The Clinical Neuropsychologist*, 23(6), 1050–1066. <https://doi.org/10.1080/13854040802665790>
- McGuire, C., Crawford, S., & Evans, J. J. (2019). Effort testing in dementia assessment: A systematic review. *Archives of Clinical Neuropsychology*, 34(1), 114–131. <https://doi.org/10.1093/arclin/acy012>
- 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. British Psychological Society.
- McWhirter, L., Ritchie, C. W., Stone, J., & Carson, A. (2020). Performance validity test failure in clinical populations: A systematic review. *Journal of Neurology, Neurosurgery & Psychiatry*, 91(9), 945–952. <https://doi.org/10.1136/jnnp-2020-323776>
- Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. University of Minnesota Press.
- Merckelbach, H., & Dandachi-FitzGerald, B. (2021). Symptomvaliditeitstests in de rechtszaal [Symptom validity tests in the courtroom]. *Expertise & Recht*, 3, 105–111.
- Merckelbach, H., Dandachi-FitzGerald, B., van Mulken, P., Ponds, R., & Niesten, E. (2015). Exaggerating psychopathology produces residual effects that are resistant to corrective feedback: An experimental demonstration. *Applied Neuropsychology: Adult*, 22(1), 16–22. <https://doi.org/10.1080/23279095.2013.816850>
- Merckelbach, H., Dandachi-FitzGerald, B., van Helvoort, D., Jelacic, M., & Otgaar, H. (2019). When patients overreport symptoms: More than just malingering. *Current Directions in Psychological Science*, 28(3), 321–326. <https://doi.org/10.1177/0963721419837681>
- Merckelbach, H., Langeland, W., de Vries, G., & Draijer, N. (2014a). Symptom overreporting obscures the dose–response relationship between trauma severity and symptoms. *Psychiatric Research*, 217(3), 215–219. <https://doi.org/10.1016/j.psychres.2014.03.018>
- Merckelbach, H., Otgaar, H., van Kan, N., & van Impelen, A. (2014b). De Schretlen Malingering Scale (MgS) als maat voor onderpresteren. [The Schretlen Malingering Scale (MgS) as measure for underperformance]. *Tijdschrift Voor Neuropsychologie*, 9(3), 239–249.
- Merten, T. (2017). Logical paradoxes and paradoxical constellations in medicolegal assessment. *Psychological Injury and Law*, 10(3), 264–273. <https://doi.org/10.1007/s12207-017-9297-3>
- Merten, T., Dandachi-FitzGerald, B., Hall, V., Schmand, B. A., Santamaría, P., & González-Ordí, H. (2013). Symptom validity assessment in European countries: Development and state of the art. *Clínica y Salud*, 24(3), 129–138. <https://doi.org/10.5093/c12013a14>
- Merten, T., & Giger, P. (2018). Wie häufig treten Simulation und Aggravation in der Begutachtung auf? Schätzungen von Laien [Lay persons' prevalence estimates of malingering in independent medical and psychological examinations]. *Forensische Psychiatrie, Psychologie, Kriminologie*, 12(2), 164–171. <https://doi.org/10.1007/s11757-018-0465-4>
- Merten, T., Kaminski, A., & Pfeiffer, W. (2020). Prevalence of overreporting on symptom validity tests in a large sample of psychosomatic rehabilitation inpatients. *The Clinical Neuropsychologist*, 34(5), 1004–1024. <https://doi.org/10.1080/13854046.2019.1694073>
- Merten, T., & Merckelbach, H. (2013). Symptom validity testing in somatoform and dissociative disorders: A critical review. *Psychological Injury and Law*, 6(2), 122–137. <https://doi.org/10.1007/s12207-013-9155-x>
- Merten, T., Merckelbach, H., Giger, P., & Stevens, A. (2016). The self-report symptom inventory (SRSI): A new instrument for the assessment of distorted symptom endorsement. *Psychological Injury and Law*, 9(2), 102–111. <https://doi.org/10.1007/s12207-016-9257-3>
- Meyer, S. R. A. (2020). *Visual associative learning in Alzheimer's Disease and performance validity: New applications of the Visual Association Test*. Doctoral thesis. Maastricht University. <https://doi.org/10.26481/dis.20200911sm>
- Meyer, S. R. A., de Jonghe, J. F. M., Schmand, B., & Ponds, R. W. H. M. (2017). The visual association test–extended: A cross-sectional study of the performance validity measures. *The Clinical Neuropsychologist*, 31(4), 798–813. <https://doi.org/10.1080/13854046.2017.1280181>
- 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(8), 1094–1102. <https://doi.org/10.1076/jcen.24.8.1094.8379>
- Mogge, N. L., & LePage, J. P. (2004). The assessment of depression inventory (ADI): A new instrument used to measure depression and to detect honesty of response. *Depression and Anxiety*, 20(3), 107–113. <https://doi.org/10.1002/da.20033>
- Monaro, M., Toncini, A., Ferracuti, S., Tessari, G., Vaccaro, M. G., De Fazio, P., & Sartori, G. (2018). The detection of malingering: A new tool to identify made-up depression. *Frontiers in Psychiatry*, 9, 249. <https://doi.org/10.3389/fpsy.2018.00249>
- Moore, P., Bunnage, M., Kemp, S., Dorris, L., & Baker, G. (2021). *Guidance on the assessment of performance validity in neuropsychological assessment*. British Psychological Society.
- Mucci, F., Avella, M. T., & Marazziti, D. (2018). Attention deficit hyperactivity disorder in adulthood: A controversial topic. *Clinical Neuropsychiatry*, 15(6), 338–343.
- Nederlands Instituut voor Psychologen, sectie Neuropsychologie. (2016). *Richtlijn neuropsychologische expertise*. [Guideline forensic neuropsychological assessment]. Retrieved from: [https://www.psynip.nl/wp-content/uploads/2016/10/RichtlijnNPexpertise\\_definitief\\_didb.pdf](https://www.psynip.nl/wp-content/uploads/2016/10/RichtlijnNPexpertise_definitief_didb.pdf)
- Nederlandse Vereniging voor Psychiatrie. (2012). *Richtlijn psychiatrisch onderzoek en rapportage in strafzaken*. [Guideline psychiatric assessment and reporting in criminal cases]. De Tijdstroom uitgeverij B.V.
- Niesten, I. J. M. (2019). *Symptom over-reporting ≠ malingering: from faulty archetypes to a nuanced empirical perspective*. Ridderprint BV. <https://doi.org/10.26481/dis.20191011in>
- Niesten, I. J. M., Müller, W., Merckelbach, H., Dandachi-FitzGerald, B., & Jelacic, M. (2017). Moral reminders do not reduce symptom over-reporting tendencies. *Psychological Injury and Law*, 10(4), 368–384. <https://doi.org/10.1007/s12207-017-9303-9>
- Orrù, G., Mazza, C., Monaro, M., Ferracuti, S., Sartori, G., & Roma, P. (2021). The development of a short version of the SIMS using machine learning to detect feigning in forensic assessment. *Psychological Injury and Law*, 14(1), 46–57. <https://doi.org/10.1007/s12207-020-09389-4>
- Pace, G., Orrù, G., Monaro, M., Gnoato, F., Vitaliani, R., Boone, K. B., & Sartori, G. (2019). Malingering detection of cognitive impairment with the b test is boosted using machine learning. *Frontiers in Psychology*, 10, 1650. <https://doi.org/10.3389/fpsyg.2019.01650>
- Pankratz, L. (1979). Symptom validity testing and symptom retraining: Procedures for the assessment and treatment of functional sensory deficits. *Journal of Consulting and Clinical Psychology*, 47(2), 409–410.
- Plohmann, A. M., & Hurter, M. (2017). Prevalence of poor effort and malingered neurocognitive dysfunction in a litigating sample in Switzerland. *Zeitschrift Für Neuropsychologie*, 28(2), 97–117. <https://doi.org/10.1024/1016-264X/a000200>
- Poloni, N., Caselli, I., Gasparini, A., Ceccon, F., Lucca, G., Ielmini, M., Isella, C., & Callegari, C. (2019). Factitious disorder as a differential diagnosis for organic hallucinations. *Minerva Psichiatrica*, 60(1), 60–64. <https://doi.org/10.23736/s0391-1772.19.02002-8>
- Pompili, M., Mancinelli, I., Girardi, P., & Tatarelli, R. (2003). Nursing schizophrenic patients who are at risk of suicide. *Journal of*

- Psychiatric and Mental Health Nursing*, 10(5), 622–624. <https://doi.org/10.1046/j.1365-2850.2003.00658.x>
- Puente-López, E., Pina, D., Ruiz-Hernández, J. A., & Llor-Esteban, B. (2021). Diagnostic accuracy of the structured inventory of malingered symptomatology (SIMS) in motor vehicle accident patients. *The Journal of Forensic Psychiatry & Psychology*, 32(1), 131–154. <https://doi.org/10.1080/14789949.2020.1833073>
- 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.
- Roma, P., Giromini, L., Burla, F., Ferracuti, S., Viglione, D. J., & Mazza, C. (2020). Ecological validity of the inventory of problems-29 (IOP-29): An Italian study of court-ordered, psychological injury evaluations using the structured inventory of malingered symptomatology (SIMS) as criterion variable. *Psychological Injury and Law*, 13(1), 57–65. <https://doi.org/10.1007/s12207-019-09368-4>
- Roma, P., Verrocchio, M. C., Mazza, C., Marchetti, D., Burla, F., Cinti, M. E., & Ferracuti, S. (2018). Could time detect a faking-good attitude? A study with the MMPI-2-RF. *Frontiers in Psychology*, 9, 1064. <https://doi.org/10.3389/fpsyg.2018.01064>
- Roor, J., Dandachi-FitzGerald, B., Peters, M., Knoop, H., & Ponds, R. (2021). Performance validity and outcome of cognitive behavior therapy in patients with chronic fatigue syndrome. *Journal of the International Neuropsychological Society*. <https://doi.org/10.1017/S1355617721000643>
- Rosen, G. M., & Phillips, W. R. (2004). A cautionary lesson from simulated patients. *Journal of the American Academy of Psychiatry and the Law*, 32(2), 132–133.
- Schlicht, D., & Merten, T. (2014). Das Bild vorgetäuschter Gesundheitsstörungen in der öffentlichen Meinung [The picture of malingered symptom presentation in public opinion]. *Versicherungsmedizin*, 66(3), 146–152.
- Schmidt, T., Watzke, S., Lanquillon, S., & Stieglitz, R. -D. (2019). *Structured interview of reported symptoms-2*. Manual. Bern, Switzerland: Hogrefe.
- Schretlen, D., Van Gorp, W. G., Wilkins, S. S., & Bobholz, J. H. (1992). Cross-validation of a psychological test-battery to detect faked insanity. *Psychological Assessment*, 4(1), 77–83.
- Sherman, E. M. S., Slick, D. J., & Iverson, G. L. (2020). Multidimensional malingering criteria for neuropsychological assessment: A 20-year update of the malingered neuropsychological dysfunction criteria. *Archives of Clinical Neuropsychology*, 35(6), 735–764. <https://doi.org/10.1093/arclin/acia019>
- Smith, G. P., & Burger, G. K. (1997). Detection of malingering: Validation of the structured inventory of malingered symptomatology (SIMS). *Journal of American Academy Psychiatry and the Law*, 25(2), 183–189.
- Stevens, A., Schmidt, D., & Hautzinger, M. (2018). Major depression – a study on the validity of clinicians' diagnoses in medicolegal assessment. *The Journal of Forensic Psychiatry & Psychology*, 29(5), 794–809. <https://doi.org/10.1080/14789949.2018.1477974>
- Stone, J., & Sharpe, M. (2020). Functional neurological symptom disorder (conversion disorder). In J. R. Geddes, N. C. Andreasen, & G. M. Goodwin (Eds.), *The new Oxford textbook of psychiatry* (3rd ed., pp. 1350–1359). Oxford University Press.
- Strubreither, W. (2021). Klinisch-neuropsychologische Untersuchung und Begutachtung. In R. Frank, W. Oder, & W. Titze (Eds.), *Das Gutachten in der gesetzlichen Unfallversicherung [Expert reports in workers' compensation contexts]* (pp. 250–282). Manz'sche Verlags und Universitätsbuchhandlung.
- Suesse, M., Wong, V. W., Stamper, L. L., Carpenter, K. N., & Scott, R. B. (2015). Evaluating the clinical utility of the medical symptom validity test (MSVT): A clinical series. *The Clinical Neuropsychologist*, 29(2), 214–231. <https://doi.org/10.1080/13854046.2015.1022226>
- SVNP. (2011). *Leitlinien für die neuropsychologische Begutachtung [Guidelines for independent neuropsychological examinations]*. [https://www.neuropsy.ch/download/pictures/d5/a1icvvhgda96j3gnrfxg5r1q5wupl9/leitlinien\\_fuer\\_die\\_neuropsychologische\\_begutachtung.pdf](https://www.neuropsy.ch/download/pictures/d5/a1icvvhgda96j3gnrfxg5r1q5wupl9/leitlinien_fuer_die_neuropsychologische_begutachtung.pdf)
- Sweet, J. J., Heilbronnner, R. L., Morgan, J. E., Larrabee, G. J., Rohling, M. L., Boone, K. B., Kirkwood, M. W., Schroeder, R. W., Suhr, J. A., & Participants, C. (2021). American Academy of Clinical Neuropsychology (AACN) 2021 consensus statement on validity assessment: Update of the 2009 AACN consensus conference statement on neuropsychological assessment of effort, response bias, and malingering. *The Clinical Neuropsychologist*. <https://doi.org/10.1080/13854046.2021.1896036>
- Van der Heide, D. H. (2021). *On the assessment of symptom validity in refugee mental health*. Doctoral thesis. Maastricht University. <https://doi.org/10.26481/dis.20210527dh>
- Van Impelen, A. (2018). *All that looks grave is not grievous. Not all those who wince are in pain: studies in furtherance of validity assessment*. Doctoral thesis. Maastricht University. <https://doi.org/10.26481/dis.20181025ai>
- Van Leeuwen, M., & de Jonghe, J. (2018). *ADI-NL Depressielijst [ADI-NL Depression Inventory]*. Hogrefe Publishing B.V.
- Viglione, D. J., & Giromini, L. (2020). *Inventory of problems-29*. Professional Manual. IOP-Test, LLC.
- Vilar-López, R., Daugherty, J., Pérez-García, M., & Piñón-Blanco, A. (2021). A pilot study on the adequacy of the TOMM in detecting invalid performance in patients with substance use disorders. *Journal of Clinical and Experimental Neuropsychology*. <https://doi.org/10.1080/13803395.2021.1912298>
- Vossler-Thies, E., Stevens, A., Engel, R. R., & Licha, C. (2013). Erfassung negativer Antwortverzerrungen mit der deutschen Fassung des "Personality Assessment Inventory", dem "Verhaltens- und Erlebensinventar." *Diagnostica*, 59(2), 73–85. <https://doi.org/10.1026/0012-1924/a000086>
- Walter, F., Petermann, F., & Kobelt, A. (2016). Erfassung von negativen Antwortverzerrungen – Entwicklung und Validierung des Beschwerdvalidierungstests BEVA [Assessment of negative response bias: Development and validation of the BEVA]. *Rehabilitation*, 55(3), 182–190. <https://doi.org/10.1055/s-0042-105939>
- Zago, S., Inglese, S., & Castiglioni, R. (2013). Malingered second-language deficit subsequent to mild traumatic brain injury. *Journal of Forensic Psychology Practice*, 13(4), 326–340. <https://doi.org/10.1080/15228932.2013.817885>
- Zago, S., Piacquadio, E., Monaro, M., Orrù, G., Sampaolo, E., Difonzo, T., & Heinzl, E. (2019). The detection of malingered amnesia: An approach involving multiple strategies in a mock crime. *Frontiers in Psychiatry*, 10, 424. <https://doi.org/10.3389/fpsyg.2019.00424>
- Zubera, A., Raza, M., Holaday, E., & Aggarwal, R. (2015). Screening for malingering in the emergency department. *Academic Psychiatry*, 39(2), 233–234. <https://doi.org/10.1007/s40596-014-0253-1>