Humanities Studios: Vol. 13, No. 4, 2025

Pedagogy, Psychology, Philosophy

Journal homepage: https://humstudios.com.ua/en

UDC 159.9:351.74:343.98

DOI: 10.31548/hspedagog/4.2025.77

Reformation of psychological assessment for firearm eligibility: Issues of diagnostics and evaluation in Austria's gun law

Salvatore Giacomuzzi*

PhD, Associate Professor

Poltava V.G. Korolenko National Pedagogical University

36000, 2 Ostrogradski Str., Poltava, Ukraine

https://orcid.org/0000-0002-4218-1685

Lisa Knoll

Master of Science

Independent researcher, Vienna, Austria

https://orcid.org/0009-0006-4626-3860

Markus Ertl

Magister

Department of Child and Youth Welfare, City of Innsbruck

6020, 5 Ing.-Etzel Str., Innsbruck, Austria

https://orcid.org/0000-0002-6408-5045

Abstract. The psychological assessment of firearm applicants is crucial for public safety, yet the scientific validity of the methods mandated by law in Austria and other European nations remains inadequate. The study emphasised the need to redefine approaches to the psychological assessment of individuals applying for firearm ownership toward greater scientific validity and interdisciplinary coherence. The study aimed to critically evaluate the diagnostic and evaluation foundations of the assessment framework enacted by the Austrian Weapons Act, highlighting scientific and legal deficiencies. The evaluation was conducted through a critical analysis of the legal regulations, a formal-logical deconstruction of the mandated test combinations, and a systematic review of the psychometric properties and forensic suitability of each prescribed evaluation instrument. The problematic legal framework, which mandated outdated and scientifically unsubstantiated test series, was analysed. The logical inconsistencies and false equivalences created by the legally permissible combinations of different psychological tests were described. The study stated that the individual instruments are unsuitable for forensic assessment due to outdated norms, a lack of predictive validity for firearm-related risk behaviour, and susceptibility to response distortion. The study further detailed how the reliance on static, actuarial assessment methods contradict modern, evidence-based standards in

Received 30.05.2025 Revised 24.10.2025 Accepted 27.11.2025

Suggested Citation:

Giacomuzzi, S., Knoll, L., & Ertl, M. (2025). Reformation of psychological assessment for firearm eligibility: Issues of diagnostics and evaluation in Austria's gun law. *Humanities Studios: Pedagogy, Psychology, Philosophy*, 13(4), 77-93. doi: 10.31548/hspedagog/4.2025.77.

*Corresponding author

forensic psychology. The results of this research can be used by legislators, legal experts, and clinical psychologists to reform the existing assessment procedures, thereby enhancing the scientific accuracy and legal defensibility of firearms licensing processes

Keywords: firearms licensing; forensic evaluation; psychometric deficiencies; forensic psychology; diagnostic validity

Introduction

Clinical psychology is no longer a subordinate auxiliary discipline but a substantive contributor to domains requiring high-stakes decision-making, especially where individual risk profiles intersect with legal and public safety concerns. One such critical area is the psychological assessment mandated by the Austrian Weapons Act (Waffengesetz, WaffG) (RIS, n.d.-b) for any individual applying for a weapon ownership license. This evaluation aims to establish an applicant's "reliability" (Verlässlichkeit), ensuring low probability of reckless firearm handling, particularly under psychological stress.

However, the scientific and legal foundation of this process is fraught with profound deficiencies that undermine the intended purpose (RIS, n.d.-b). The significance of robust licensing procedures is underscored by public health research. Such studies as K. Kennedy *et al*. (2025) and S. Oliphant *et al*. (2025), from the United States, for instance, consistently demonstrate a significant association between the benefits of firearm purchaser licensing laws and reductions in firearm-related deaths, including those among young people and in the context of intimate partner violence. While the legal frameworks differ, these findings highlight the high stakes involved in worldwide firearms legislation. Therefore, the psychological assessments at the core of such laws should be scientifically valid, reliable, and legally defensible – a standard the Austrian system, as this article demonstrated, fails to meet. The diagnostic framework imposed by law is rooted in outdated practices and relies on assessment tools that are inconsistent, psychometrically weak, and ill-suited for the complex task of prognostic evaluation.

The core of the problem lies in the legal requirement for psychologists to use a fixed, or "taxative", list of tests, a practice that directly contradicts the principle of scientific progress and professional autonomy. In a comprehensive critique, A. Kampusch (2022) argued that the term "reliability" in the 1996 firearms act is insufficiently defined, rendering it scientifically impossible to operationalise and verify. Following the study results, this ambiguity creates systematic distortions in the significance of any test used. The author further identified major flaws in the mandated procedures, including the use of outdated validation instruments, redundancies between test subscales, and the fact that only three of the fifteen listed tests include validity scales to detect response distortion. Compounding this issue, a study by M. Ertl *et al*. (2021) empirically demonstrated that most test subjects were able to identify and manipulate these few validity scales, therefore inefficient at cheating prevention.

The Austrian assessment model is further compromised by its adherence to a static, actuarial paradigm, a methodology that has been largely superseded in modern forensic psychology. International best practice has shifted decisively toward dynamic, context-sensitive risk assessment, often referred to as structured professional judgment (SPJ). For instance, seminal work by D. Brookstein *et al*. (2020) on the HCR-20^{V3} have established the superior predictive validity of models that integrate historical, clinical, and future-oriented dynamic risk factors. In their comprehensive international handbook, K. Douglas & R. Otto (2021) confirmed that contemporary risk assessment

acknowledges the limitations of static instruments, emphasising that individual behaviour cannot be captured solely through cross-sectional test results. This perspective is reinforced by T. Albright & N. Scurich (2024), highlighting in the study on psychopathy and legal decision-making the necessity of using multi-faceted assessment tools that can account for complex personality dynamics, a feature absent in the rigid Austrian framework. This undermines large-scale meta-analyses that have shown that structured instruments significantly outperform static or unstructured evaluation in predictive accuracy. Furthermore, the European psychiatric association has explicitly underpinned the value of an evidence-based, dynamic and context-sensitive approach as the professional standard (Völlm *et al.*, 2018).

This legislative and scientific inertia is not unique to Austria but reflects a broader European challenge. Modern studies show that similar issues persist across the continent. In Germany, for example, the legal guidelines for psychological examinations remain vague, leading to significant inconsistencies in evaluation practices between different jurisdictions (Hoppe *et al.*, 2020). Similarly, in Italy, the lack of a nationally binding assessment protocol results in significant disparities ("heterogeneous solutions") in the quality and methodology of assessments between regions, determining the scientific need for greater uniformity (Blandino *et al.*, 2024). In Poland, critical analyses of the legal framework highlight significant discrepancy between legal definitions and forensic science, confirming a lack of standardised quality control for the required psychological examinations. These international examples underscore the urgent need for a paradigm shift towards an evidence-based, scientifically robust approach. The continued use of outdated, static models not only fails to meet modern professional standards but also creates a significant legal paradox, holding experts liable for reports based on scientifically indefensible methods.

The study aimed to critically analyse the diagnostic and test-theoretical foundations of the psychological assessments for firearm reliability legally prescribed in Austria. By deconstructing the mandated test batteries and their underlying assumptions, significant scientific and legal inconsistencies were revealed. The analysis first examined the problematic legal framework, then demonstrated the logical fallacies of the prescribed test combinations through a formal analysis and concluded with a detailed critique of the suitability of each individual test instrument, based on the official test manuals and provider documentation (Testzentrale, n.d.; Schuhfried, n.d.; Pearson, n.d.). Lastly, the far-reaching implications of using outdated, static assessment models were discussed, and a paradigm shift toward a modern, evidence-based approach was proposed.

Issues of the firearms law implementation and the test procedures prescribed therein

Section 3 paragraph 2 of the first firearms act implementation (1. WaffV) regulates which test procedures are to be used during the assessment for determining reliability (RIS, n.d.-a). To ensure meaningful assessments at a scientifically recognised level, so-called "established test procedures" were already selected in 1997 and listed (taxativ) in the 1. WaffV in coordination with various "experts". With the further amendment of the 1. WaffV in 2012, the Federal Minister of the Interior expanded the existing selection with various test combinations that were considered qualitatively and "scientifically equivalent". Following the requirement of 8 paragraph 7 WaffG, the procedures were listed "taxatively" (i.e., fully, exhaustively) (cf. 3 paragraph 2a 1. WaffV). However, since that time, no further development of diagnostics or verification was amended to cover modern scientific research and expertise. Studies on the issue in the context of Austria are insufficient.

In this context, it is worth noting from a scientific-professional perspective that the term "taxative" mentioned in the legal text cannot be valid, as knowledge is never complete and a

normative determination may perpetuate an "error or only preliminary finding", which contradicts practical developments. Accordingly, diagnostic recommendations can never be "taxative", especially when they are outdated and partly obsolete from a professional standpoint. In this context, the ongoing turn away from actuarial, static, and statistical diagnostic instruments, where the field has progressed to the fifth developmental stage of prognostic and idiographic assessments, should be referenced.

D. Gerhold & C. Obermaier (2013) stated that, similar to the previously familiar approach, experts were still "free" to use the base scales of the Minnesota multiphasic personality inventory 2 (MMPI-2 base scales) in combination with the stress coping questionnaire (Stressverarbeitungsfragebogen, S-V-F) or the reliability-oriented personality test – version 3 (Verlässlichkeitsbezogener Persönlichkeitstest, VPT-3), alongside risk propensity factors questionnaire (Fragebogen für Risikobereitschaftsfaktoren, F-R-F). These two test series largely correspond to those already familiar and well-established. The legal text was updated to list the then-prevailing versions of the procedures. D. Gerhold & C. Obermaier (2013) list about 20 combinations of assessment instruments that would theoretically be possible under the law.

According to the amendment of the 1st weapons act implementation ordinance (1. Waffengesetz-Durchführungsverordnung, 1. WaffV) in 2012, several new combinations of test instruments were introduced as alternatives to the previously used combination of the MMPI-2 base scales with the stress processing questionnaire (Stressverarbeitungsfragebogen, S-V-F) (RIS, n.d.-a). These new regulations stipulated that the Eppendorf schizophrenia scale (Eppendorfer Schizophrenie-Inventar, ESI) must be employed. This was to be combined with the short questionnaire for aggression factors (Kurzfragebogen zur Erfassung von Aggressivitätsfaktoren, K-FAF) and the locus of control questionnaire (Fragebogen zu Kontrollüberzeugungen, IPC). Additionally, one of several stress assessment tools was required, providing the choice between the S-V-F, the coping inventory for stressful situations (CISS), or the differential stress inventory HR (Differentielles Stressinventar HR, DSIHR).

To complete the assessment, examiners were to choose one of four distinct pathways. The first option was to add the personality style and disorder inventory (Persönlichkeits-Stil- und Störungs-Inventar, PSSI). A second alternative was to use the inventory of clinical personality accentuations (Inventar Klinischer Persönlichkeitsakzentuierungen, IKP) instead of the PSSI. The third pathway involved adding the NEO five-factor inventory (NEO-Fünf-Faktoren-Inventar, NEO-FFI) along with a depression screening tool, which could be either the general depression scale (Allgemeine Depressionsskala, ADS) or the beck depression inventory revision (BDI-II). The fourth and final option was to use the big five plus one personality inventory (Big Five Plus One Persönlichkeitsinventar, B5PO), also supplemented by either the ADS or the BDI-II. At the time, the Federal Minister of the Interior, according to D. Gerhold & C. Obermaier (2013), wanted to state using the term "or" that the test procedures named could be combined as an alternative. Thus, experts "would have" the possibility to choose from the above-mentioned combinations and prescribe them. This would also significantly hinder the occurrence of a "learning effect" in candidates when undergoing repeated weapons-psychological evaluations.

A. Kampusch (2022) shows in a study that the direct influence of the 1st weapons act implementing ordinance (1. WaffV) on the selection options of clinical-psychological test procedures applied in phase 1 evaluations in Austria creates systematic errors. The study criticised the fact that the term "reliability" in the 1996 firearms act (WaffG) is insufficiently defined and thus cannot be scientifically operationalised and verified. This circumstance causes systematic distortions in the test results. Simulation of desirable traits as well as dissimulation of undesirable traits also occur

among applicants. The combination of the test instruments leads to redundancies between the various subscales, the test instruments are partially outdated and/or have deficiencies in the normative samples. Moreover, A. Kampusch (2022) criticised that only three of the 15 listed test procedures have validity scales to identify tendencies toward response distortion.

Additionally, the study by M. Ertl *et al.* (2021) demonstrated that most test subjects were able to manipulate the validity scales, therefore they are insufficiently suitable for identification of response distortion. The listed procedures are already internally inconsistent, from both an expert and professional-scientific point of view, and contradictory in their specific constructs, original goals, areas of application, and psychometric properties – so that the presumed "prognostic and content equivalence", especially among all these possible combinations, is not present. Moreover, the combinatory permutation itself is flawed, difficult to comprehend, and has never been tested regarding its predictive validity for lawful behaviour.

The problematic pseudo-validity of the construct "reliability" in the firearms act implementation

The Austrian legislation assumes that the test instruments explicitly listed in the implementation ordinance can be combined interchangeably and are all equally suitable for assessing the construct of "reliability" as required in paragraph 8 of the firearms act (WaffG). For analytical clarity, the construct "reliability" (regardless of what it may represent in psychological reality) will hereafter be abbreviated as V. According to the legal text, the construct of Reliability (V) may be inferred through various specific combinations of diagnostic instruments. At least formally, this results in a series of supposed equivalences, where V is assessed by the following combinations:

Combination 1:
- Minnesota multiphasic personality inventory-2 (MMPI-2)
- Stress coping questionnaire (SVF)

Combination 2:
- Minnesota multiphasic personality inventory-2 (MMPI-2)
- Risk propensity questionnaire (FRF)
- Reliability-oriented personality test (VPT-3)

Combination 3:
- Eppendorf schizophrenia inventory (ESI)
- Short questionnaire for aggression factors (K-FAF)
- Locus of control questionnaire (IPC)
- Stress coping questionnaire (SVF)
- Personality style and disorder inventory (PSSI)

Combination 4:
- Eppendorf schizophrenia inventory (ESI)
- Short questionnaire for aggression factors (K-FAF)
- Locus of control questionnaire (IPC)
- Coping inventory for stressful situations (CISS)
- Personality style and disorder inventory (PSSI)

Combination 5:
- Eppendorf schizophrenia inventory (ESI)
- Short questionnaire for aggression factors (K-FAF)
- Locus of control questionnaire (IPC)
- Differential stress inventory HR (DSIHR)
- Personality style and disorder inventory (PSSI)

Combination 6:
- Eppendorf schizophrenia inventory (ESI)
- Short questionnaire for aggression factors (K-FAF)
- Locus of control questionnaire (IPC)
- Stress coping questionnaire (SVF)
- Inventory of clinical personality accentuations (IKP)

Combination 7:
- Eppendorf schizophrenia inventory (ESI)
- Short questionnaire for aggression factors (K-FAF)
- Locus of control questionnaire (IPC)
- Coping inventory for stressful situations (CISS)
- Inventory of clinical personality accentuations (IKP)

Combination 8:
- Eppendorf schizophrenia inventory (ESI)

- Short questionnaire for aggression factors (K-FAF)
- Locus of control questionnaire (IPC)
- Differential stress inventory HR (DSIHR)
- Inventory of clinical personality accentuations (IKP)

...etc.

It is possible to deduce a set of logically inconsistent identifies from these combinations through basic algebraic transformation. For example: the MMPI-2 is treated as equivalent to the combined use of the ESI, the K-FAF, the IPC, and the PSSI. Furthermore, the stress coping questionnaire (SVF) is interchangeable with the combination of the risk propensity questionnaire (FRF) and the reliability-oriented personality test (VPT-3). The SVF is also treated as equivalent to both the CISS and the DSIHR. Similarly, the PSSI is considered diagnostically interchangeable with the inventory of IKP.

From these legally implied equivalences, it is possible to incorrectly deduce several scientifically indefensible conclusions: this implies that three psychometrically distinct instruments – the SVF, the DSIHR, and CISS – are functionally identical. The regulation also assumes that the comprehensive Minnesota multiphasic personality inventory-2 (MMPI-2) is equivalent to the sum of four other tests (ESI, K-FAF, IPC, and PSSI), and that the inventory of IKP and the PSSI are interchangeable. Furthermore, it is suggested that the SVF can be directly substituted by combining the FRF with the reliability-oriented personality test (VPT-3). Consequently, this combination of the FRF and VPT-3 is also considered equivalent to both the DSIHR and the CISS.

These formulaic derivations arithmetically suggest that different diagnostic instruments are functionally equivalent in their capacity to assess V, which contradicts not only basic psychometric principles but also the theoretical and clinical foundations of the instruments involved. Moreover, such equivalences were never intended by the original test developers and defy construct validity and contextual fit. For instance, it is methodologically and diagnostically indefensible to assume that the MMPI-2 is functionally equivalent to the sum of ESI, K-FAF, IPC, and PSSI. These instruments differ fundamentally in their diagnostic focus, empirical foundation, and psychometric properties.

The inconsistencies become even more pronounced considering the full range of theoretically permissible combinations – of which over 20 exist under the regulation. Each new equation yields further contradictions that render the diagnostic logic arbitrary and incoherent. From a formal-logic standpoint, these identities are invalid: from a false premise, no valid conclusion can follow. These theoretical concerns have been empirically substantiated. A. Kampusch (2022) evaluated 100 tests collected from firearms psychological assessments between 2015 and 2021, which included combinations such as SVF-120, ESI, K-FAF, IKP, and IPC. The findings were unequivocal: all such combinations exhibited significant tendencies toward response distortion and systematic redundancy. All tested inventories demonstrated high intercorrelations among subscales, implying substantial overlap in content and construct representation. This redundancy causes statistical inflation and undermines the validity and reliability of any conclusions regarding V (Doyle & Dolan, 2002). Especially concerning is that many of the instruments are outdated or were never intended for forensic or prognostic application.

Furthermore, the observation by A. Kampusch (2022) that only 3 out of the 15 mandated instruments include validity or manipulation-check scales (e.g., lie scales) further weakens the framework. These concerns are compounded by the findings of M. Ertl *et al.* (2021), demonstrating that most test-takers were able to recognise and strategically circumvent such scales, further reducing their utility in detection of simulation or dissimulation. In sum, the regulatory assumption that heterogeneous instruments can be freely combined while still maintaining content and prognostic equivalence is deeply flawed. The result is a system in which logically contradictory test combinations are mandated yet treated as

diagnostically and legally interchangeable. No professional or empirical justification for this approach can be found in the international literature. The regulatory framework thereby constitutes a diagnostic and legislative paradox that undermines the scientific credibility of firearms psychological assessments in Austria.

Critical review of the proposed test instruments from a diagnostic and test-theoretical perspective

For the sake of clarity and academic rigor, the test instruments listed in the Austrian firearms act implementation (1. WaffV, paragraph 3) are presented and systematically examined for their diagnostic validity, test-theoretical robustness, and forensic applicability (RIS, n.d.-a). The scientific suitability of the instruments for assessment of legal construct of "reliability" (Verlässlichkeit), as stipulated by the Austrian firearms law, was emphasised (Table 1, Fig. 1). The ADS represents a German-language screening tool adapted from the centre for epidemiological studies depression scale (CES-D) (Giacomuzzi & Ertl, 2024). While its utility in public health screening is acknowledged, its methodological limitations, such as limited normative samples and insufficient psychometric standardisation, render it unsuitable for forensic purposes. The absence of predictive validation concerning firearm risk further questions its applicability. The BDI-II is widely utilised in clinical settings to assess the severity of depressive symptoms. However, critiques noted outdated normative data and problematic German translations (Margraf & Schneider, 2018). No empirical studies support its predictive value for firearm-related risk behaviours, nor is it constructed to assess stress tolerance or impulse control – key components in determining legal reliability (Heilbrun *et al.*, 2014). The B5PO operationalises broad personality traits but lacks fine-grained clinical sensitivity. Although based on the five-factor model, its abbreviated item structure limits its discriminant capacity (John *et al.*, 2008). Its utility for high-stakes legal evaluations is therefore limited, as it was never normed for forensic populations.

Table 1. *Available test instruments by scales and item count*

Test name	Scales	Items	Provider
MMPI-2 – Minnesota multiphasic personality inventory®-2, basic scales	13	370	Testzentrale (Hogrefe)
ESI – Eppendorfer schizophrenia inventory	4	40	Testzentrale (Hogrefe)
K-FAF – short questionnaire on aggressiveness factors	5	49	Testzentrale (Hogrefe)
IPC – questionnaire on locus of control	3	24	Testzentrale (Hogrefe)
S-V-F – stress coping questionnaire	19	114	Testzentrale (Hogrefe)
CISS – coping inventory for stressful situations	3	48	Pearsonclinical (UK)
DSIHR – differential stress inventory	13	121	Schuhfried GmbH
PSSI – personality style and disorder inventory	14	140	Testzentrale (Hogrefe)
IKP – inventory of clinical personality accentuations	11	110	Testzentrale (Hogrefe)
NEO-FFI – NEO five-factor inventory	5	60	Testzentrale (Hogrefe)
B5PO – big five plus one personality inventory	6	66	Schuhfried GmbH
ADS – general depression scale	1	20	Testzentrale (Hogrefe)
BDI-II – beck depression inventory revision	1	21	Testzentrale (Hogrefe) / Schuhfried GmbH
VPT.3 – reliability-oriented personality test	6	70	Schuhfried GmbH
FRF – risk propensity questionnaire	3	49	Schuhfried GmbH
FPI-R – Freiburg personality inventory	12	138	Testzentrale (Hogrefe)

Source: *compiled based on the official test catalogues and manuals of the listed providers, including Testzentrale (n.d.), Schuhfried (n.d.), Pearson (n.d.)*

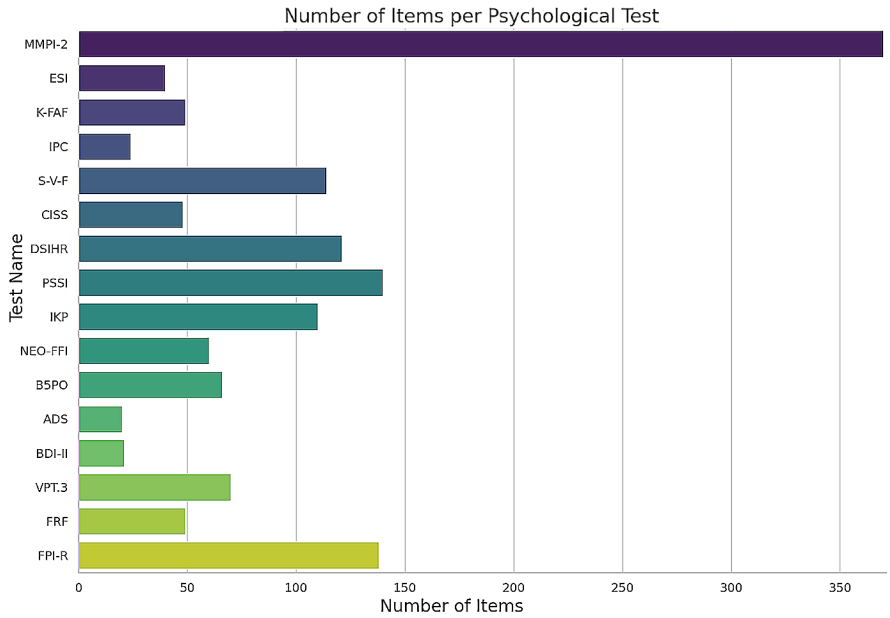

Figure 1. *Number of items per psychological test*

Note: *MMPI-2: Minnesota multiphasic personality inventory-2; ESI: Eppendorf schizophrenia inventory; K-FAF: short questionnaire on aggressiveness factors; IPC: questionnaire on locus of control; S-V-F: stress coping questionnaire; CISS: coping inventory for stressful situations; DSIHR: differential stress inventory; PSSI: personality style and disorder inventory; IKP: inventory of clinical personality accentuations; NEO-FFI: NEO five-factor inventory; B5PO: big five plus one personality inventory; ADS: general depression scale; BDI-II: beck depression inventory revision; VPT.3: reliability-oriented personality test; FRF: risk propensity questionnaire; FPI-R: Freiburg personality inventory*

Source: *compiled by the authors based on data in Table 1*

The CISS measures habitual coping styles across three dimensions: task-oriented, emotion-oriented, and avoidance-oriented coping. Although it demonstrates solid psychometric qualities in general populations, its relevance to forensic risk assessments remains tenuous (Endler & Parker, 1990; Andrade, 2009). Coping behaviour in experimental conditions may not translate to real-world firearm safety scenarios. O. Kohut (2020) noted that the Differential Stress Inventory (DSI) evaluates stress responses but lacks transparency regarding its theoretical underpinnings and discriminant validity. Empirical support remains limited, and the distinction between DSI and DSI-HR is not sufficiently clarified (Schuhfried, n.d.). The construct's tangential connection to firearm misuse or recklessness undermines its inclusion in reliability diagnostics. The ESI has demonstrated utility in screening for schizophrenia-spectrum disorders. Nevertheless, the presence of a mental disorder such as schizophrenia is neither necessary nor sufficient for establishing ineligibility in legal contexts. Notably, many individuals with psychotic disorders demonstrate stable behaviour under treatment (APA, 2022).

In a study by M. Ertl *et al.* (2021), 60% of subjects successfully manipulated the Freiburg personality inventory – revised (FPI-R) social desirability scales. Earlier studies had already identified its vulnerabilities to response bias. Without effective validity scales, the instrument is not capable of reliably detecting dissimulation or simulation of real-world behaviour. The FPI-R should therefore not be used in isolation for legal assessments of firearms eligibility. The FRF measures physical, social, and financial risk-taking. However, its

limited external validity in predicting real-life risk behaviour, especially in forensic settings, renders its use highly questionable. No studies link FRF scores with impulsive aggression or firearm misuse. The IKP offers structured assessment of personality disorders and accentuations. Although methodologically rich, its forensic applicability is limited by the absence of operational thresholds for determining reliability. Furthermore, empirical links between specific personality disorder profiles and firearm misuse are scarce (Paris, 2003).

The IPC differentiates between internal and external control beliefs. However, substantial concerns regarding its factorial structure, outdated normative data, and interpretability remain. Its role in predicting behaviour under duress is speculative at best. Forensic practice demands updated tools with validated control indices (Roesch & Cook, 2017; Ortega-Escobar *et al.*, 2017). The K-FAF aims to detect trait aggression. While normed, it still lacks comprehensive validation studies regarding falsification and forensic utility. No studies confirm that aggression scores are predictive of unsafe firearm use (Hutten *et al.*, 2024). Despite extensive clinical use, the MMPI-2 is limited in forensic validation in its German version. Several subscales lack adequate reliability and predictive validity. The MMPI-2-RF, a more modern revision, offers improved psychometric properties but is not yet standard in Austrian forensic psychology (Ben-Porath, 2023). This gap between the mandated procedures and current scientific standards is further highlighted by research based on the even more modern MMPI-3 for firearm injury prevention studies (Bandel *et al.*, 2025). No version has been validated for predicting legal reliability.

The NEO-FFI captures key personality dimensions. Although its theoretical foundation is substantial, the German version lacks necessary psychometric metrics for forensic use (Costa & McCrae, 1992). Furthermore, it offers insufficient insight into impulsivity, aggression, or disinhibition – core traits in firearm-related risk. The PSSI, rooted in PSI-theory, offers a phenomenological model of personality disorders. Its small sample sizes and low validity coefficients limit its forensic relevance. As no studies correlate PSSI dimensions with legal constructs of reliability, its inclusion in firearm assessments is questionable. The SVF differentiates coping strategies but demonstrates limited validity in predicting behaviour under stress. Studies indicate strong situational modification of coping patterns (Trempa *et al.*, 2002). The predictive power of the SVF-120 is too weak for use in legal risk prognoses. Reliability-based personality test – version 3 (VPT.3). The VPT.3 is insufficiently documented, with no independent validation studies (Schuhfried, n.d.). It lacks transparency in scoring, theoretical grounding, and test fairness. As of 2024, the test is not indexed in international psychological databases nor referenced in any peer-reviewed forensic journals.

The diagnostic heterogeneity of the test battery specified in the 1. WaffV, along with its lack of conceptual alignment with the legal construct of "reliability", renders its forensic use methodologically unsound (RIS, n.d.-a). Many instruments are outdated, poorly normed, or lack any predictive value for firearm misuse or risk behaviour. More fundamentally, no empirical study exists to justify the arbitrary combination of such diverse tests into a coherent diagnostic battery. From a scientific perspective, the substitution or permutation of these instruments, some with 370 items, others with less than 50, cannot yield consistent or valid results. The continued reliance on these outdated methods contradicts best practices in forensic psychology and raises serious concerns regarding the legal and ethical robustness of the firearms licensing process in Austria.

On the problematic use of purely statistical and static assessment instruments

The goal of a reliability prognosis is to predict the psychological stability of individuals regarding potential future behaviour – essentially in assessment of the probability of hidden adverse events.

Even the most performant prognostic models lose validity in two years. Furthermore, internationally recognised diagnostic systems such as the ICD-10/ICD-11 or DSM-5-TR were never designed to generate prognostic predictions regarding "reliability" and cannot validly represent this construct (APA, 2022; WHO, n.d.).

In addition to the already mentioned problematic permutations, it is necessary to critically assess the psychometric quality of the instruments used. Of all the instruments listed in the Austrian ordinance, only three include validity scales (e.g., lie scales or openness indices) intended to detect socially desirable responding – MMPI-2, FPI-R, and ESI. However, these scales are controversial due to limited reliability in detecting cheating (Ben-Porath, 2023).

For instruments that lack such validity indices, the probability of response bias increases significantly, especially as many of them are easily interpretable by test-savvy individuals. Further analysis of the issue demonstrates the several tests are below an ROC (receiver operating characteristic) curve threshold of 0.7 – indicating insufficient discriminative accuracy. The ROC curve plots true positive rate (sensitivity) against false positive rate (1 – specificity), and an AUC (area under the curve) below 0.7 suggests limited predictive utility (Swets, 1988; Ellerbrock, 2010; Youngstrom, 2014). Modern studies highlighted these limitations: F. van Leeuwen *et al.* (2025) underpinned that AUC values of prognostic model change significantly across external validation studies. This implies that large uncertainty when applying an instrument to a new context. Furthermore, R. Denz *et al.* (2023) highlighted that AUC is often misused as an all-purpose performance metric in risk assessments.

A. Kampusch (2022) highlighted that the test procedures mandated by the 1. WaffV are not only partially redundant at the subscale level but also outdated, no longer commercially available in some cases, and based on obsolete or insufficiently documented normative samples. These limitations fundamentally undermine the fairness and validity of any comparative conclusions drawn from such data. Modern forensic risk assessment uses AUC values derived from ROC analyses to establish the predictive validity of instruments. State-of-the-art tools such as the HCR-20^{V3} or the FoVOx rely on a combination of historical, clinical, and future-oriented dynamic risk factors – reflecting a shift toward structured professional judgment and evidence-based idiographic evaluation (Douglas *et al.*, 2014).

Statistical or static diagnostic instruments, such as those often cited in the Austrian context, are therefore further defined as obsolete. Contemporary risk prediction acknowledges that individual trajectories cannot be captured solely through historical data or cross-sectional test results. Instead, dynamic interactional models – considering context, psychosocial variability, and clinical judgment – have proven more accurate (Ogloff, 2003; Koh & Kim, 2023). Moreover, some case constellations simply defy conventional diagnostic categorisation. For these, specialised instruments such as FOTRES or FoVOx are recommended, though their application requires structured training and calibration (Fazel *et al.*, 2012; Goncalves *et al.*, 2017). Unfortunately, such sophisticated, dynamic approaches remain absent from Austrian firearms reliability assessments, which still rely on standardised self-report tests with fixed answer categories. None of the instruments in use possess scientifically verified predictive validity for the complex construct of behavioural reliability under stress, social provocation, or legal scrutiny.

In summary, the continued reliance on outdated, largely static, and insufficiently validated instruments – absent a theoretical foundation linking them to the legal construct of "reliability" – undermines the credibility of the assessment framework. To ensure legal robustness and scientific rigor, a paradigm shift toward idiographic, evidence-based, and dynamically grounded diagnostics is urgently needed. This necessity mirrors a broader, international demand for greater scientific accountability in all forensic disciplines, including a call for more transparency, open data,

and rigorous validation of methods, often summarised under the principle of "open science" (Albright & Scurich, 2024). Adopting these principles is fundamental to establishing the scientific and legal integrity of future reliability assessments.

Legal issues and conclusion of observations

The foregoing analysis demonstrated that the assessment procedures mandated under the Austrian firearms act (WaffG) for evaluation of "reliability" is not scientific or robust (RIS, n.d.-b). The entire framework is based on a foundation of outdated science, logical fallacies, and legislative inertia, creating a significant legal and professional paradox for the experts tasked with its implementation. The central issue is the legal obligation for experts to use a fixed, "taxative" list of diagnostic instruments that are, as demonstrated, scientifically indefensible. The construct of "reliability" that these tests are meant to assess remains insufficiently defined within the law, rendering it impossible to operationalise and validate in a scientifically rigorous way. This fundamental ambiguity means that any subsequent assessment, regardless of the instruments used, rests on a legally precarious foundation. A judgment of "unreliability" derived from such an undefined premise is highly vulnerable to legal challenge, as it lacks objective, verifiable criteria. Experts are therefore required to produce legally binding judgments on a construct that has no clear scientific basis, using tools that are ill-suited for the task. This stands in evident contrast to modern forensic practice, which demands that legal constructs have clear, empirical referents to be assessed validly (Ortega-Escobar *et al.*, 2017).

This creates a direct and irreconcilable conflict with their professional and ethical obligations. According to paragraph 22. 1 of the psychologists' act, experts are guaranteed "freedom in professional practice", which includes autonomy over the choice of diagnostic instruments. Furthermore, they are legally bound by professional standards to align their practice with the current state of psychological science. However, the state of the science has moved decisively away from the static, actuarial instruments listed in the ordinance. International best practice is centred on dynamic, evidence-based models of SPJ, such as the HCR-20^{V3}, which integrate historical, clinical, and dynamic future-oriented risk factors and have far superior predictive validity (Douglas *et al.*, 2014; Pirelli *et al.*, 2015). The exclusive reliance on outdated, item-based self-report instruments no longer reflects scientific standards and fundamentally contradicts the principles of evidence-based practice in forensic psychology.

The legal framework forces a system where psychometrically distinct instruments are treated as functionally interchangeable, leading to logically contradictory and arbitrary assessment batteries. Yet experts remain personally liable for their reports. They are caught in an untenable position: adhere to a flawed law and violate their professional duties, or adhere to professional standards and violate the letter of the weapons act. This flawed process not only exposes experts to significant liability, including potential civil lawsuits from applicants or sanctions from professional bodies, but it also raises serious questions about due process for applicants. Their fundamental eligibility for a firearms license is decided by scientifically indefensible methods, which is therefore impossible to appeal a negative decision on objective, evidence-based grounds (Heilbrun *et al.*, 2014). The contradiction in the law is further highlighted by the fact that in phase 2 assessments, experts are granted full discretion in their choice of tools, an inconsistency that undermines the entire regulatory logic. This legislative and scientific inertia is not an isolated Austrian anomaly but reflects a broader European challenge.

The Austrian case in an international context

The problems delineated throughout this article, ranging from outdated diagnostic instruments and a reliance on static-actuarial methods to the logically inconsistent combination of tests, are

not exclusive to Austria. While the Austrian case is particularly notable for its rigid, legally mandated list of flawed test combinations, the underlying issues of legislative inertia and a failure to align legal requirements with modern scientific standards are echoed across the continent. A comparative analysis of practices in other European nations reveals a widespread and troubling inconsistency in how psychological fitness for firearm ownership is assessed, highlighting the urgent need for a paradigm shift toward an evidence-based, scientifically robust approach.

Several countries, such as Germany and Italy, provide a substantial contrast to Austria's rigid framework. In Germany, a psychological examination may be required under paragraph 6 of the Waffengesetz (WaffG), but the legal guidelines remain vague, leaving the selection of instruments to the examiner's discretion. While this lack of a binding federal standard creates significant inconsistencies in evaluation practices, it avoids the specific legal and ethical paradox found in Austria. German experts are not legally compelled to use a fixed list of scientifically indefensible tools; the Austrian expert, however, is bound by a "taxative" list that mandates outdated procedures and creates a direct conflict with their professional duty to use state-of-the-art methods.

A similar situation is present in Italy, where assessments are handled regionally without a nationally binding protocol, leading to significant disparities in quality and methodology. Although some Italian experts use instruments, such as the MMPI-2, they are not forced to treat it as diagnostically equivalent to a collection of other, completely different tests – a core logical fallacy of the Austrian system. An Italian expert could, in theory, apply a modern, evidence-based approach like SPJ, whereas the Austrian expert is legally barred from doing so in the critical phase 1 assessment. The problem in these countries is a lack of standardisation, leading to inconsistent quality; Austria's problem is the standardisation of scientifically flawed practices. Other nations, such as the UK and Nordic countries, have even less formalised systems. In England and Wales, psychological assessment is not a routine requirement and is typically only initiated if "good reason" suggests a potential issue, with the evaluation of "fitness" often left to general practitioners and police forces.

Similarly, in Sweden and Norway, psychological assessment is rarely formalised and tends to be reactive rather than proactive. A study by S. Henriksen & B. Kruke (2020) confirmed that these assessments vary widely and lack structured prognostic models. While this approach forgoes the potential benefit of systematic screening, it also avoids the systemic harm of mandating a scientifically invalid process for every applicant. The issue in these jurisdictions can be seen as a potential under-utilisation of psychological expertise, which starkly contrasts with Austria's legally enforced misapplication of it.

The challenges of Poland highlight shared methodological deficiencies. In Poland, psychological assessments are mandated, but analyses show insufficient quality control regarding the validity, predictive utility, or critically – the manipulation resistance of the tests being used. This directly mirrors the Austrian dilemma, where only three of the fifteen mandated instruments even include validity scales to detect response distortion, and those have been empirically shown to be largely ineffective. The reliance on tests that cannot reliably detect faking is a profound forensic shortcoming shared by both systems.

These international cases reveal a clear and global inconsistency in the psychological assessment of firearm applicants. The widespread reliance on static self-report tools, outdated diagnostic logic, and the absence of standardised, empirically validated methods reflects a profound forensic shortcoming across Europe. This comparative analysis strengthens the conclusion that the Austrian model is not just flawed but represents a rigid and paradoxical failure of regulation. The urgent need for a paradigm shift towards scientifically sound, evidence-based standards is not just an Austrian issue, but an international imperative for ensuring public safety.

Conclusions

The analysis of the psychological assessment of firearm reliability in Austria, presented in the study, causes several critical conclusions. First, the legal framework established by the 1st weapons act implementation ordinance is scientifically untenable. The "taxative" or exhaustive list of procedures effectively freezes the assessment process in time, ignoring decades of scientific progress. It compels experts to use test procedures that are outdated, questionable from a test-theoretical standpoint, and unsuitable for the prognostic assessment of the construct "reliability – a term the law fails to adequately define, rendering any scientific operationalisation impossible from the outset. Second, the assumption that the various prescribed test combinations are diagnostically equivalent and interchangeable is a logical fallacy. As has been shown, this leads to formally and logically inconsistent equations of psychometrically distinct instruments, such as treating the comprehensive MMPI-2 as equal to a combination of four entirely different tests. This arbitrary interchangeability undermines the validity of the entire assessment process from a basic scientific perspective. Third, the individual test instruments lack suitability for the forensic context. Many suffer from outdated norms, low predictive validity for risk-relevant behaviour, and high susceptibility to socially desirable responding. This vulnerability is compounded by the fact that few possess effective validity scales, and even those have been empirically shown to be easily circumvented by informed test-takers, rendering them largely ineffective in high-stakes evaluations. Fourth, anchoring the assessment exclusively in a static-actuarial paradigm contradicts scientific standards in forensic risk prognosis. Contemporary, evidence-based approaches call for a shift towards dynamic methods such as SPJ, which integrate historical, clinical, and future-oriented risk factors – methods absent in the Austrian system.

In summary, the system places experts in an irreconcilable legal and ethical dilemma, caught between the flawed mandates of the weapons act and their professional duty under the psychologists act to adhere to the state of the science. This fundamentally challenges the scientific credibility of firearm-related assessments in Austria. A fundamental reform, guided by scientific findings, is urgently needed to ensure the rule of law, professional integrity, and public safety. Prospects for future research should therefore focus on the validation of dynamic, evidence-based risk assessment tools within the Austrian legal context and on developing a scientifically operationalised model of "reliability" to inform future legislation.

Acknowledgements

Not applicable.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

[1] Albright, T.D., & Scurich, N. (2024). A call for open science in forensics. *Proceedings of the National Academy of Sciences of the United States of America*, 121(24), article number e2321809121. doi: 10.1073/pnas.2321809121.

[2] American Psychiatric Association (APA). (2022). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington: American Psychiatric Publishing. doi: 10.1176/appi.books.9780890425787.

[3] Andrade, J.T. (2009). *Handbook of violence risk assessment and treatment: New approaches for forensic mental health professionals*. New York: Springer.

[4] Bandel, S.L., Rodriguez, T.R., Anestis, M.D., & Anestis, J.C. (2025). Utilizing personality assessment to further firearm injury prevention research: An example investigation with the MMPI-3. *Journal of Personality Assessment*, 1-9. doi: 10.1080/00223891.2025.2550381.

[5] Ben-Porath, Y.S. (2023). *Interpreting the MMPI-2-RF*. Minneapolis: University of Minnesota Press.
[6] Blandino, A., Galante, N., Cuppone, F., Giriodi, M., & Travaini, G.V. (2024). Firearm ownership and suicide: Has the time come to discuss uniformity of health and social assessments in aid of regulation? Reflection from a retrospective study on a forensic case series. *Annali di Igiene*, 37(2), 215-224. doi: 10.7416/ai.2024.2648.
[7] Brookstein, D.M., Daffern, M., Ogloff, J.R.P., Campbell, R.E., & Chu, C.M. (2020). Predictive validity of the HCR-20^{V3} in a sample of Australian forensic psychiatric patients. *Psychiatry, Psychology, and Law*, 28(3), 325-342. doi: 10.1080/13218719.2020.1775152.
[8] Costa, P., & McCrae, R.R. (1992). Neo PI-R professional manual. *The NEO Inventories*, 6, 244-256.
[9] Denz, R., Klaaßen-Mielke, R., & Timmesfeld, N. (2023). A comparison of different methods to adjust survival curves for confounders. *Statistics in Medicine,* 42(10), 1461-1479. doi: 10.1002/sim.9681.
[10] Douglas, K.S., & Otto, R.K. (2021). *Handbook of violence risk assessment* (2nd ed.). New York: Routledge.
[11] Douglas, K.S., Hart, S.D., Webster, C.D., Belfrage, H., Guy, L.S., & Wilson, C.M. (2014). Historical-clinical-risk management-20, version 3 (HCR-20^{V3}): Development and overview. *The International Journal of Forensic Mental Health,* 13(2), 93-108. doi: 10.1080/14999013.2014.906519.
[12] Doyle, M., & Dolan, M. (2002). Violence risk assessment: Combining actuarial and clinical information to structure clinical judgements for the formulation and management of risk. *Journal of Psychiatric and Mental Health Nursing*, 9(6), 649-657. doi: 10.1046/j.1365-2850.2002.00535.x.
[13] Ellerbrock, D. (2010). Gun violence and control in Germany 1880-1911: Scandalizing gun violence and changing perceptions as preconditions for firearm control. In W. Heitmeyer, H.G. Haupt, S. Malthaner & A. Kirschner (Eds.), *Control of violence* (pp. 185-212). New York: Springer. doi: 10.1007/978-1-4419-0383-9_8.
[14] Endler, N.S., & Parker, J.D. (1990). Multidimensional assessment of coping: A critical evaluation. *Journal of Personality and Social Psychology*, 58(5), 844-854. doi: 10.1037/0022-3514.58.5.844.
[15] Ertl, M., Giacomuzzi, S., Gundolf, E.-M., & Garber, K. (2021). Evaluation of social desirability (lie scales) using the freiburg personality inventory (FPI-R) especially in regard to child custody procedures. *Psychologie in Österreich,* 2, 169-176.
[16] Fazel, S., Singh, J.P., Doll, H., & Grann, M. (2012). Use of risk assessment instruments to predict violence and antisocial behaviour in 73 samples involving 24 827 people: Systematic review and meta-analysis. *BMJ,* 345, article number e4692. doi: 10.1136/bmj.e4692.
[17] Gerhold, D., & Obermaier, C. (2013). Reliability under firearms legislation in Austria. *SIAK-Journal – Zeitschrift für Polizeiwissenschaft und Polizeiliche Praxis*, 3, 47-59. doi: 10.7396/2013_3_E.
[18] Giacomuzzi, S., & Ertl, M. (2024). On the fundamental problems of diagnostics and test theory in psychological assessments for the purpose of firearms legislation. *Psychologie in Österreich,* 4 & 5, 336-345.
[19] Goncalves, L.C., Rossegger, A., & Endrass, J. (2017). Forensic operationalized therapy/risk evaluation system (FOTRES). In J.P. Singh, D.G. Kroner, J.S. Wormith, S.L. Desmarais & Z. Hamilton (Eds.), *Handbook of recidivism risk/needs assessment tools* (1st ed.) (pp. 243-253). Hoboken: Wiley.
[20] Heilbrun, K., DeMatteo, D., Holliday, S.B., & LaDuke, C. (Eds.). (2014). *Forensic mental health assessment: A casebook* (2nd ed.). Oxford: Oxford University Press. doi: 10.1093/med:psych/9780199941551.001.0001.
[21] Henriksen, S.V., & Kruke, B.I. (2020). Norwegian police use of firearms: Critical decision-making in dynamic and stressful situations. *Nordic Journal of Studies in Policing*, 7(2), 99-120. doi: 10.18261/issn.2703-7045-2020-02-03.

[22] Hoppe, S.J., Zhang, Y., Hayes, B.E., & Bills, M.A. (2020). Mandatory arrest for domestic violence and repeat offending: A meta-analysis. *Aggression and Violent Behavior*, 53, article number 101430. doi: 10.1016/j.avb.2020.101430.

[23] Hutten, J.C., van Horn, J.E., Hoppenbrouwers, S.S., Ziermans, T.B., & Geurts, H.M. (2024). Neuropsychological assessment of aggressive offenders: A Delphi consensus study. *Frontiers in Psychology*, 15, article number 1328839. doi: 10.3389/fpsyg.2024.1328839.

[24] John, O.P., Naumann, L.P., & Soto, C.J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement and conceptual issues. In O.P. John, R.W. Robins & L.A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed.). (pp. 114-158). New York: Guilford Press.

[25] Kampusch, A. (2022). System error in the weapons psychology assessment of phase 1 by the 1st weapons act implementation ordinance. *Sachverständige*, 4, 207-211.

[26] Kennedy, K.S., Wagner, E.D., Meyerson, N.S., Oliphant, S.N., McCourt, A.D., Zeoli, A.M., Webster, D.W., & Crifasi, C.K. (2025). Examining the impacts of firearm purchaser licensing laws on firearm deaths among youth aged 15 to 24, by age group-USA, 1990-2019. *Injury Prevention*, 0, 1-5. doi: 10.1136/ip-2025-045700.

[27] Koh, M.K., & Kim, H.S. (2023). Discriminant analysis of high-risk recidivism in criminal offenders based on psychopathological factors from MMPI-2-RF profiles. *Journal of Forensic Psychiatry & Psychology,* 34(3-4), 332-356. doi: 10.1080/14789949.2023.2226126.

[28] Kohut, O. (2020). The concept of integrative development of the personality's tolerance to stress in systemic discourse. *Scientific Bulletin of Mukachevo State University. Series "Pedagogy and Psychology"*, 6(2), 113-123. doi: 10.52534/msu-pp.6(2).2020.113-123.

[29] Margraf, J., & Schneider, S. (2018). *Textbook of behavioural therapy* (4th ed.). Heidelberg: Springer Berlin.

[30] Ogloff, J.R.P. (2003). Advances in violence risk assessment: A structured professional judgment approach. *Australian and New Zealand Journal of Psychiatry*, 37, article number A5-A5.

[31] Oliphant, S.N., Crifasi, C.K., Doucette, M.L., Topazian, R.J., McCourt, A.D., Kennedy, K.S., & Zeoli, A.M. (2025). Association of firearm purchaser licensing laws with intimate partner homicide. *Violence and Gender*, 12(2), 67-73. doi: 10.1089/vio.2024.0035.

[32] Ortega-Escobar, J., Alcázar-Córcoles, M.A., Puente-Rodríguez, L., & Peñaranda-Ramos, E. (2017). Psychopathy: Legal and neuroscientific aspects. *Anuario de Psicología Jurídica*, 27(1), 57-66. doi: 10.1016/j.apj.2017.01.003.

[33] Paris, J. (2003). Personality disorders over time: Precursors, course, and outcome. *Journal of Personality Disorders*, 17(6). doi: 10.1521/pedi.17.6.479.25360.

[34] Pearson. (n.d.). Retrieved from https://www.pearsonclinical.de/.

[35] Pirelli, G., Wechsler, H., & Cramer, R.J. (2015). Psychological evaluations for firearm ownership: Legal foundations, practice considerations, and a conceptual framework. *Professional Psychology: Research and Practice,* 46(4), 250-257. doi: 10.1037/pro0000023.

[36] RIS. (n.d.-a). *Consolidated federal law: Complete legislation on firearms 1996 (1. WaffV)*. Retrieved from https://surl.li/kyfyph.

[37] RIS. (n.d.-b). *Consolidated federal law: Complete legislation on firearms 1996 (1. WaffV)*. Retrieved from https://surl.li/yswyfh.

[38] Roesch, R., & Cook, A.N. (2017). *Handbook of forensic mental health services*. New York: Routledge. doi: 10.4324/9781315627823.

[39] Schuhfried. (n.d.). Retrieved from https://www.schuhfried.com/.

[40] Swets, J.A. (1988). Measuring the accuracy of diagnostic systems. *Science*, 240(4857), 1285-1293. doi: 10.1126/science.3287615.

[41] Testzentrale. (n.d.). Retrieved from https://www.testzentrale.de/.

[42] Trempa, E., Janke, W., & Weyers, P. (2002). The predictability of stress processing methods under natural stress conditions using the SVF120. *Zeitschrift für Differentielle und Diagnostische Psychologie,* 23(3). doi: 10.1024/0170-1789.23.3.291.

[43] van Leeuwen, F.D., Steyerberg, E.W., van Klaveren, D., Wessler, B., Kent, D.M., & van Zwet, E.W. (2025). Instability of the AUROC of clinical prediction models. *Statistics in Medicine,* 44(5), article number e70011. doi: 10.1002/sim.70011.

[44] Völlm, B.A., Clarke, M., Herrando, V.T., Seppänen, A.O., Gosek, P., Heitzman, J., & Bulten, E. (2018). European Psychiatric Association (EPA) guidance on forensic psychiatry: Evidence-based assessment and treatment of mentally disordered offenders. *European Psychiatry*, 51, 58-73. doi: 10.1016/j.eurpsy.2017.12.007.

[45] World Health Organization (WHO). (n.d.). *International classification of diseases 11th revision. The global standard for diagnostic health information*. Retrieved from https://icd.who.int/.

[46] Youngstrom, E.A. (2014). A primer on receiver operating characteristic analysis and diagnostic efficiency statistics for pediatric psychology: We are ready to ROC. *Journal of Pediatric Psychology*, 39(2), 204-221. doi: 10.1093/jpepsy/jst062.

Реформа психологічної оцінки для отримання права на володіння вогнепальною зброєю: питання діагностики та оцінки у законодавстві Австрії про зброю

Сальваторе Джакоміцці
Кандидат філософських наук, доцент
Полтавський національний педагогічний університет імені В.Г. Короленка
36000, вул. Остроградського, 2, м. Полтава, Україна
https://orcid.org/0000-0002-4218-1685
Ліза Кнолл
Магістр наук
Незалежний дослідник, м. Відень, Австрія
https://orcid.org/0009-0006-4626-3860
Маркус Ертль
Магістр
Департамент у справах дітей та молоді
6020, вул. Інн.-Етцеля, 5, м. Інсбрук, Австрія
https://orcid.org/0000-0002-6408-5045

Анотація. Психологічна оцінка осіб, які подали заяву на отримання дозволу на володіння вогнепальною зброєю, має вирішальне значення для громадської безпеки. Дослідження підкреслило необхідність переосмислення підходів до психологічної оцінки осіб, які претендували на право володіння зброєю, у напрямі більшої наукової валідності та міждисциплінарної узгодженості. Водночас наукова обґрунтованість методів, передбачених законодавством Австрії та інших європейських країн, залишається вкрай сумнівною. Метою цієї статті було здійснити критичну оцінку діагностичних і тестово-теоретичних засад системи оцінювання, встановленої законом Австрії про зброю та виявлення її наукових і юридичних недоліків. Оцінювання здійснювалося шляхом критичного аналізу нормативно-правових положень, формально-логічної деконструкції обов'язкових комбінацій тестів та систематичного огляду психометричних характеристик і судово-психологічної придатності кожного із передбачених інструментів. У статті проаналізовано проблемну правову базу, яка вимагає використання застарілих і науково необґрунтованих тестових батарей. Описано логічні неузгодженості та хибні еквіваленти, що виникають унаслідок законодавчо допустимих комбінацій різних психологічних тестів. Зазначено, що окремі інструменти здебільшого непридатні для судово-психологічної оцінки через застарілі норми, відсутність прогностичної валідності щодо ризику поведінки, пов'язаної зі зброєю, та вразливість до спотворення відповідей. Детально показано, що виключна залежність від статичних, актуарних методів оцінювання суперечить сучасним, доказовим стандартам судової психології. Результати цього дослідження можуть бути використані законодавцями, юристами та клінічними психологами для реформування чинних процедур оцінювання, що сприятиме підвищенню наукової обґрунтованості та юридичної захищеності процесу видачі дозволів на володіння зброєю

Ключові слова: дозвіл на володіння зброєю; судово-психологічна оцінка; психометричні недоліки; судова психологія; діагностична валідність