Chapter # 25

RISK APPRAISAL AND LEGAL PRINCIPLES
Unveiling disciplinary gaps

Martina Feldhammer-Kahr[1,2], Nina Kaiser[2], Ida Leibetseder[2], Markus Sommer[1], & Martin E. Arendasy[1]
[1]Department of Psychology, University of Graz, Austria
[2]Department of Criminal Law, Criminal Procedure Law and Criminology, University of Graz, Austria

ABSTRACT
This chapter strives to enhance the understanding of the challenges of risk assessment in a legal context while building bridges between the disciplines. Forensic psychology aims, upon other things, to assess recidivism risk. The Austrian jurisprudence also focuses on aspects of behavioral prognosis: the verdict should aim at preventing recidivism. Thus, judges must follow provisions that emphasize (current) factors related to the individual offender. This requires the interplay of different scientific perspectives and a strong interdisciplinary practice. Therefore, psychological risk appraisal guides are used to provide valid indicators for recidivism risks. However, the lack of substantive discussion on expert opinions in court, combined with the high frequency of courts adopting them verbatim in verdicts, bears risks in this rather multidisciplinary than interdisciplinary practice. A closer look at the German versions of the VRAG-R (Rettenberger, Hertz, & Eher , 2017) and the LSI-R (Dahle, Harwardt, & Schneider-Njepel, 2012), reveals that some aspects and their unquestioned application require critical legal reflection. Furthermore, the HCR-20V3 (Müller-Isberner, Schmidbauer, & Born, 2014) reveals weaknesses in the practical risk assessment, potentially leading to similar problems, if misapplied. This chapter focuses on these problem areas: (1) quality of expert opinions, (2) individuality and topicality, (3) legal reality.

Keywords: violent risk assessment, risk instruments, legal principle, interdisciplinarity.

1. INTRODUCTION

In the Austrian justice system, especially in Austrian criminal law, there is a focus on behavioral predictions, driven by a preventive justice approach. Sentences aim to deter individual offenders from reoffending (specific deterrence) and to prevent the general population from committing crimes while strengthening the public's awareness of the legal consequences of criminal behavior and the enforcement of these consequences (general deterrence). According to prevailing opinions, criminal judgments should primarily serve specific deterrent purposes (cf. Jerabek & Ropper, 2024; Kaiser, 2023; Schroll & Oshidari, 2024). This leads to a focus on offender-related factors and ultimately to the need of an individual-centered prognosis, that the law alone cannot provide (Kaiser, 2023; Kaiser & Leibetseder, 2024). Consistently, Austrian judges emphasize the need to integrate psychological, criminological and pedagogical factors to adequately assess the individual's future behavior (Leibetseder, Kaiser, & Woschizka, 2024). Thus, judicial officials need to consult and integrate non-legal disciplines (Kastner, 2019). In Austria, judges and prosecutors are obliged to appoint expert witnesses whenever special expertise is required for investigations or the taking of evidence, which the prosecuting authorities do not have at their disposal through their bodies, special institutions or persons permanently employed by

them (§ 126 Austrian Code of Criminal Procedure). This is generally done by appointing expert witnesses who provide their professional opinions on the grounds of their findings through thorough exploration and application of the proper tests (see: Kastner, 2019). However, the integration of "foreign" disciplines in legal proceedings with the aim of seeking empirically grounded facts, brings both opportunities and new challenges (Bock, 2017; Glatz & Aicher, 2019). Every science has its own perception of entities, which leads to divergent paradigms, attitudes and premises. Hence, as the cornerstone of the legitimacy of interdisciplinary work, every integration of a "foreign" discipline, such as psychological-statistical prognostic instruments in legal proceedings, must be contextualized in a discipline-specific way.

In law, there is little corresponding discourse on the content and foundation of applied prognostic tools. Nevertheless, in criminal proceedings, it is the judge's responsibility to evaluate expert opinions provided, like any other evidence. A task which is rather difficult, because the judge needs to evaluate a report with knowledge they do not possess (cf. Kabzinska, 2021) (e.g. quality criteria of assessment tools). In practice, such an evaluation is scarcely evident during proceedings (Kastner, 2019). Fegert et al.'s (2006) examination of expert opinions in sexual offense cases in Germany (N=38) revealed that 31.6% showed blanked approval without hesitation. 10.5% of judges did not rephrase the expert's opinion and 21.1% rephrased the text from the report, both groups did not discuss the content. 10.5% rephrased the assessment report and did discuss the content. 7.9% considered the content without mentioning it. 13.2% followed the verbal report of the expert witness. Approximately 95% did follow the expert witness. This is not merely a habit in dealing with predictions for sexual offenders. This study also shows a lack of evaluation in culpability assessments, which conclude an overall adoption rate of culpability assessments for sexual offenders of 95.1%, and for homicide and arson offenses, of 88.6%. This suggests that critical reflections of expert witness reports are rather rare.

Due to the broad similarities in the legal systems, we can assume a similar situation in Austria. Judges are trained in legal matters but have not built up an expertise in psychological assessments as well as psychologists are experts in their field but are not trained in law. This bears the risk that incongruencies in the scientific approaches are overlooked. Both parties might perform their own discipline flawless without noticing the missing link of mutual reflection in the realm of an effective integrative interdisciplinary work that guarantees compatibility and validity. The danger associated with ignoring the necessary moderation of interdisciplinary cooperation will be illustrated below by examining three problem areas (aspects of quality, individuality and topicality, legal reality) arising with the application of the VRAG-R (Harris, Rice, Quinsey, & Cormier, 2015; German: Rettenberger et al., 2017), the LSI-R (Andrews & Bonta 1995; German: Dahle et al., 2012), and the HCR-20^{V3} (Douglas, Hart, Webster, & Belfrage, 2013; German: Müller-Isberner, Schmidbauer, & Born, 2014), three of the most common actuarial risk assessment instruments international (cf. Neal & Grisso, 2014) and in Austrian correctional facilities.

2. ASPECTS OF QUALITY

According to several studies, one of the most important indicators for the selection of expert witnesses through judges is the expert's work experience (cf. Tadei, Finnilä, Korkman, Salo, & Santtila, 2014; Tadei, Finnilä, Reite, Antfolk, & Santtila, 2016), although the work experience in the clinical field does not necessarily lead to more valid judgements of the professionals (cf. Garb, 1989). A study of Kabzinska (2021) in Poland showed that judges prefer expert witnesses with whom they already worked or who were recommended. Also,

sometimes there are no other expert witnesses which could be appointed simply due to the lack of available experts (cf. Kastner, 2019).

Another problem which occurs is a potential preselection of evaluators and the results. According to studies, expert witnesses tend to (consciously or unconsciously) reach conclusions that support the party who retained them (Guarnera, Murrier, & Boccaccini, 2017). Results from structured risk instruments showed higher scores, and therefore a higher expected risk of recidivism, in cases of prosecution retained evaluations, while defense-retained scores were lower (e.g. Murrie et al., 2009). All parties involved in the system should be aware of this bias. Transparency in the evaluation process from the expert witness and a better understanding of risk assessments on the legal side would help prevent such effects.

As mentioned by Shapiro, Mixon, Jackson, and Shook (2015), judges may need more instructions on how to differentiate between reliable and unreliable expert testimony (see also: Kovera & McAuliff, 2000). The required information cannot come from within the legal community but the respective field, whereby an interdisciplinary dialog is nevertheless crucial for the integration of this knowledge in this very specific and highly sensitive context.

3. INDIVIDUALITY AND TOPICALITY

By ensuring that the punishment is primarily intended for specific deterrence, the legislator expresses its focus on the individual case and the individual offender in criminal proceedings. This brings to the fore those behavioral and personal sentencing criteria that can address and capture the offender in his social and behavioral context. This seems to be correct insofar as the results of criminological studies indicate that future criminal behavior cannot be predicted only through assessing risk factors such as “broken home” or social class. Instead, an appropriate assessment relies upon closer examination of specific social conditions (Bock, 2017; c.f. Walton, 2022). Even an overview of several tools, which is interpreted individually by the expert, does not seem to mitigate this potential conflict (Bock, 2017).

Thus, it seems perplexing that the challenge of necessary individualization in criminal proceedings is addressed most of the time with merely a nomothetic research logic, which requires establishing regularities and thus generalization, although modern psychological diagnostics recommend a combination of nomothetic and explanative approaches for the assessment (Schmidt-Atzert, Krumm, & Amelang, 2021).

Such generalization can indeed serve as additional evidence in proceedings, but it would need to be appropriately weighed by the judge, which, as the above-mentioned study shows, is hardly happening. In view of “hindsight bias”, “confirmation bias” and a multitude of other heuristics that “subjectify” human judgment (Wistrich, Rachlinski, & Guthrie, 2015; Guthrie, Rachlinski, & Wistrich, 2002), an amendatory synthesis by the respective judge cannot provide the necessary individuality (Bock, 2017). In fact, such a non-transparent and supposedly intuitive overall view in turn undermines the objectivity that the nomothetic prognosis conveys (Bock, 2017).

The current practice may therefore violate the principles of Austrian criminal law by categorizing individuals solely based on their “risk profile” (cf. Bock, 2017; Brettel, 2022; Riffel, 2023). Additionally, deciding on the necessary criminal law interventions for an individual, by mere statistical membership of a risk group, seems to neglect the fatality of alpha (false positive) errors in statistical procedures when used in legal proceedings. Can sentences be legitimized if they knowingly and deliberately accept wrong decisions in individual cases in favor of a rule-based overall view? Considering criminal law being the

ultima ratio of state authority, it seems not to be in the interest of a modern, individual, prevention-oriented sentencing practice.

When looking closer at specific types of assessment tools, these concerns regarding the lack of individuality become more concrete. One of the most used second-generation tools is the Violence Risk Appraisal Guide (VRAG-R; Quinsey, Harris, Rice, & Cormier, 2006; German: Rettenberger et al., 2017). Using statistical techniques to provide an assessment of the level of risk, based on static variables (historical risk factors), such as the number of convictions, is much better than the first generation of risk assessment, where the decision-making process is unguided (Fletcher, Gredecki, & Turner, 2022, p. 46). Studies indicate that the VRAG-R shows an overall good predictive validity for violent offenders (e.g. Rice, Harris, & Lang, 2013) and the criteria of objectivity is given (cf. Fletcher et al., 2022). But some problems still arise. Regarding the need of a person-centered prognosis, it seems concerning that neither the motive nor the psychological functioning of the offenders is taken into account (for further reading: Rossegger et al., 2011). On the other hand, there are various items which can be criticized regarding content validity (Rossegger, Gerth, & Endrass, 2013). For example, Item 1 deals with the "separation from either parent (except death) under the age of 16", which does make sense considering the test construction process, where the authors searched for variables which highly correlate with recidivism (Rossegger et al., 2013). Even so, since the control group rarely includes non-delinquent individuals, it is questionable whether these correlations can have explanatory potential in predicting the risk of delinquency (Bock, 2017). Nonetheless, we would have to look on the specific case, because it could still be preferable to be separated from a violent parent or a toxic environment. Using such indicators without taking a closer look to the particular case may present a problem with the criteria fairness and objectivity.

Another fundamental principle of a preventive justice approach focused on specific deterrence is that the sentence and its imposed interventions are based on present circumstances (topicality). It is essential to ensure a prognosis that is sensitive to changes, responding appropriately to potential alterations in criminologically relevant factors, remaining dynamic, and not solely focusing on aspects related to the past. Looking again at the VRAG-R, one major problem is that there are mainly static variables included and changes in the life of the offender do not affect the score and therefore the predicted level of risk (cf. Rossegger et al., 2013). If the level of change is not measured the offender has no possibility to get a lower estimated risk and thus the assessment tool does not meet the requirements of a preventive justice approach. The remark in the additional explanations to the manual, that in the case of an offense-free period in freedom, the recidivism rate can be optionally reduced by 10%, does not only strongly depend on the assessor, but makes it even harder for the judge to independently question the statistical results regarding their topicality (Rettenberger et al., 2017). This indicates that recourse to the VRAG-R under the guise of specific deterrence therefore not only lacks the principle of individuality, but also the principle of topicality.

On the contrary, the so called third generation risk assessments (e.g. Level of Service Inventory–Revised: LSI-R; Andrews & Bonta, 1995; German: Dahle et al., 2012) include dynamic items (e.g. emotional/personal, companions, alcohol/drugs) and allow the documentation of needs and changes (Rossegger et al., 2011). In considering changes in several areas the offender has the possibility to get a lower estimation of their risk of recidivism level. Although, this assessment tool seems to fit better for the purpose at hand, it includes 10 items regarding to the criminal history (e.g. arrested under age 16). Therefore, it can be hard for individuals with a long criminal record to get a low probability (< 15%) no matter what changed in their individual life.

Lastly, it should be noted, that the LSI-R score alone cannot be used to predict the recidivism (c.f. Dahle et al., 2012; Harwardt & Schneider-Njepel, 2013), nor should others. In fact, it should be used for the assessment of needs and changes. As a result of the lack of integrative interdisciplinary work such aspects are often overlooked in the context of legal proceedings. The mentioning of concrete estimated recidivism rate, on the base of the LSI-R raw score, therefore, gives a deceptive sense of security in the decision-making process of the court.

There was also concern regarding the calculation of a sum score for risk estimation from HCR-20 Version 2 (Webster, Douglas, Eaves, & Hart, 1997) users (cf. Müller-Isberner, Jöckel, & Gonzales Cabeza, 1998), because of the different item topics. The authors of the HCR-20^{V3} (Douglas et al., 2013) considered feedback of hundreds of users of Version 2 to improve the risk assessment tool. Due to the uncomfortableness with the numeric rating system the third version uses letters (no, possibly/partially, yes) to rate the presence of the item contents (e.g. H9. Violent attitudes, C2. Violent ideation or intent, R3. Personal support), instead of numbers. Also, the relevance (low, moderate, high) of the items for the specific individual are evaluated. In addition, the user is provided with a "seven-step structure to guide risk assessment and development of treatment and management strategies" (Bjørkly, Eidhammer, & Selmer, 2014, p. 235). Furthermore, information on risk management was added and a specification of how risk factors might operate uniquely for the individual being tested (for overview see: Douglas et al., 2014). Thus, the tool certainly provides promising developments in tackling the problems mentioned above.

The proper usage requires extensive knowledge in several areas. It needs file reviews, different sources of historical information, interviews with the evaluatee and other sources, observations and much more (see: Douglas et al., 2014). This is a time-consuming process. Even the authors conclude that "not all steps will be done by all evaluators or in all cases" (Doulas et al., 2014, p. 100). In doing so the quality criteria of the tool will no longer persist. That's practical reality (see: Kastner, 2019).

As can be seen, there is a lot of scientific work done to improve certain risk assessment tools and there is a lot of work ahead.

4. LEGAL REALITY

Another major point which needs to be considered is the legal reality, where adapted values of foreign law collide with the Austrian law and inherent fundamental principles.

In addition to the purpose of punishment, fundamental procedural rights suffer from an uncritical incorporation of statistical figures into the judgment. Item 5 and Item 8 of the VRAG-R address the assessment of the criminal history, drawing on the Canadian system by Cormier and Lang (1999), based on the earlier version by Akman and Normandeau (1967) for classifying violent and non-violent offenses. Offenses in the subjects' past are evaluated with a predetermined score, which is then added up. The descriptions in the items of the VRAG-R manual add that in cases where there is a discrepancy between the indictment and the conviction, the more serious of the two should be used, which will often be the indictment (Rettenberger et al., 2017). Therefore, the VRAG-R explicitly allows scoring not the actual offense the individual was convicted for in accordance with the rule of law but the suspicion and thus an initial assumption that has not (legally and/or factually) proven to be true. However, the principles of the presumption of innocence, as well as the dominant principles *nullum crimen, nulla poena sine lege* ("no crime, no penalty without law") are particularly crucial for the rule of law in penal practice. The presumption of innocence in accordance with Section 8 of the Code of Criminal Procedure and Article 6 (2) of the European Convention

on Human Rights (ECHR) is a "principle that governs the entire Austrian legal system", which means that "the measure of punishment can only be guilt established in accordance with procedural law". If the sentencing is based on an offense that was neither the subject of the proceedings nor of any other legally binding conviction, the principle is deemed to have been violated. According to the case law of the ECHR, assumptions of suspicion can in turn be compatible with the presumption of innocence under Art 6 (2) ECHR (scope of protection Art 6 (2) "until proven guilty according to law") in the case of an examination of the personality of the perpetrator carried out as part of the sentencing process (Grabenwarter, 2021). It should be noted, that in the present case it is not a suspicion that is used to examine the personality of the perpetrator as a part of the sentencing process but a charge from a previous case that has already been established as unsubstantiated, thus insufficient for a conviction. Therefore, the equal treatment of charges and convictions should not be seen as an examination of personality but as an inadequate prognosis that is assessed and is decisively considered when sentencing. According to the ECHR, this might in turn be sufficient for a violation of the presumption of innocence. Also, in view of the possibly even "stricter" Austrian legal basis ("only the guilt established in accordance with the procedure") in the interpretation of Article 6, undermining these principles using psychological tools that do not follow these premises should therefore be carefully reconsidered with respect to potential infringements. Another question that arises regards the utilization of already expunged convictions. According to the established case law of the Austrian Higher Administrative Court, the facts underlying the expunged convictions, i.e. the "behavior" of the person concerned, may be considered for the assessment of this risk prognosis in the context of aliens or residence law and also in the case of background checks under the Weapons Act, but not the expunged convictions themselves (Kert, 2017). Kert (2017) correctly sees a tension here with the presumption of innocence. According to Section 1 (2) of the Austrian Redemption Act the law provides for the effect of "full expungement", which means that all adverse consequences extinguish, and the person is considered as having no criminal record. It does not distinguish between different levels of expungement and thus should not be weakened (Kert, 2017). While, as previously discussed, the challenge often lies in judges having to weigh psychological aspects, now the challenge is that expert witnesses have to consider legal aspects when carrying out their assessments.

It is also worth reflecting on the fact that the Cormier-Lang system is based on Canadian Law. As Canada's legal system is rooted in British and French Common Law, it works differently from Austrian Law as a representative of a civil law system. Other than Austria, Canadian law is thus strongly influenced by case law. Whenever you transfer "foreign" law into your own, but especially when such diametrically opposed systems clash, you need to take a closer look. The authors of the German version of the VRAG-R point out that it is therefore not always possible to clearly classify the offenses. The offense category that "best describes the course of events" should be selected after additional surveys. At this point, it should not be forgotten that the respective case handlers are not lawyers and that such a classification can be tricky in individual cases. In addition, even if the case is handled conscientiously, it cannot be ruled out that behavior is assessed and scored as an offense in the VRAG-R that is not necessarily relevant under Austrian law. In our case the integration of Canadian Criminal Law to Austrian Criminal Law leads to (at least) three substantial challenges. (1) The Cormier-Lang system assesses behavior that is under no circumstance punishable by Austrian criminal law (e.g. "disturbance"). The explanations of the German Version state, that one exclusively has to work with the offender's official criminal record, but caution is required here for users as the VRAG-R in other places or for in-depth study allows to assess other documents, such as witness, or police reports or other information

provided by the offender. (2) There is also the risk of scoring behavior that, in individual cases, might not be punishable by Austrian criminal law. For example, “possession of items for burglary” or “wearing a mask with the intention of committing a crime” might not reach the threshold of a criminally punishable attempt stage under Austrian criminal law. In addition, further inconsistencies could arise, for example in that complicity is given a higher score, although the Austrian criminal law does not qualitatively differentiate between “role-categories” of offending. (3) Lastly, even the comparability of those offenses, which are undoubtedly punishable in Austria and in Canada, must be questioned. One must not take for granted, that the elements of the offense under Canadian law and the Austrian law are the same leading to the same indictment or conviction. Even between similar criminal law systems, different legal opinions on the same offense groups have developed, leading to inconsistencies: For example, Austria and Germany have differences in the definition and legal treatment of murder and manslaughter, even if the terms appear similar at first glance. In Austria, murder (§ 75 Austrian Criminal Code) is defined more broadly and does not require specific murder characteristics (such as malice or base motives) which is the case in Germany. This shows that an unquestioned indirect influence of the system and values of foreign law is problematic one should be aware of when applying and evaluating the results in proceedings.

5. CONCLUSION

Risk assessment puts the legal system, and all involved to the test. The focus of the article was to give readers a better understanding of the risk assessment process and problems that may occur in several areas. Even if they may not occur in one specific case, there should be an awareness about potential risks for the individual and the legal system.

One must take a closer look which parts of the decision-making process already work and in which one there is still some work to be done to bridge disciplinary divides and meet both legal and assessment criteria. In doing so, we have the opportunity to increase the fairness for the individual and meet the requirements of Austrian law. The goal can only be reached through uninterrupted interdisciplinary exchange, explicit demands for reports, and the broadening of knowledge and awareness on both sides.

REFERENCES

Akman, D. D., & Normandeau, A. (1967). The measurement of crime and delinquency in Canada. *British Journal of Criminology*, *7*(2), 125–128.

Andrews, D., & Bonta, J. (1995). *The Level of Service Inventory - Revised manual*. Toronto: Multi-Health Systems.

Bjørkly, S., Eidhammer, G., & Selmer, L. E. (2014). Concurrent validity and clinical utility of the HCR-20^{V3} compared with the HCR-20 in forensic mental health nursing: similar tools but improved method. *Journal of Forensic Nursing*, *10*(4), 234-242. https://doi.org/10.1097/JFN.0000000000000047

Bock, M. (2017). *Angewandte Kriminologie* [Applied Criminology]. Mainz: Johannes Gutenberg University.

Brettel, H. (2022). Zur Kriminologie als forensischer Disziplin. [On criminology as a forensic discipline]. *Forensische Psychiatrie, Psychologie, Kriminologie*, *16*, 277-285. https://doi.org/10.1007/s11757-022-00740-w

Cormier, C., & Lang, C. (1999). Cormier-Lang System for Quantifying Criminal History. In V.L. Quinsey, G.T. Harris, M.E. Rice, & C.A. Cormier (Eds.), *Violent Offenders – Appraising and Managing Risk* (3rd ed.). Washington, DC: American Psychological Association.

Dahle, K.-P., Harwardt, F., & Schneider-Njepel, V. (2012). *Inventar zur Einschätzung des Rückfallsrisiko und des Behandlungsbedarfs von Straftätern. Deutsche Version des Levels of Service Inventory-Revised nach Don Andrews und James Bonta (LSI-R)* [German Version of the Levels of Service Inventory-Revised from Don Andrews and James Bonta.]. Göttingen: Hogrefe.

Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). *HCR-20^{V3} Assessing Risk for Violence*. Burnaby, BC: Mental Health, Law, and Policy Institute, Simon Fraser University.

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-20V3): Development and overview. *International Journal of Forensic Mental Health*, *13*(2), 93-108. https://doi.org/10.1080/14999013.2014.906519

Fegert, J. M., Schnoor, K., König, C., & Schläfke, D. (2006) Psychiatrische Begutachtung in Sexualstrafverfahren. Eine empirische Untersuchung zur Schuldfähigkeit bei jugendlichen, heranwachsenden und erwachsenen Beschuldigten in Mecklenburg-Vorpommern. [Psychiatric assessment in sexual offense proceedings. An empirical study on the culpability of juvenile, adolescent and adult defendants in Mecklenburg-Vorpommern]. In J.M. Fegert, (Eds.) *Schriften zum Jugendrecht und zur Jugendkriminologie* (Vol.7). [Series of publications on juvenile law and juvenile criminology]. Herbolzheim: Centaurus.

Fletcher, M., Gredecki, N., & Turner, P. (2022). Forensic risk assessment. In N. Gredecki & P. Turner (Eds.), *Forensic Psychology* (pp. 44-56). New York, NY: Routledge.

Garb, H. N. (1989). Clinical judgment, clinical training, and professional experience. *Psychological Bulletin*, *105*(3), 387-396. https://doi.org/10.1037/0033-2909.105.3.387

Glatz, M., & Aicher, G. (2019). Gutachten im Spannungsfeld zwischen Beauftragung und Qualität. [Expert opinions in the field of tension between commissioning and quality.]. In G. Brinek. (Ed.), *Gutachten als Schlüsselfaktoren im Maßnahmenvollzug. Schriftenreihe der Volksanwaltschaft* [Expert witness reports as key factors in the correctional system. Series of publications from the Austrian Ombudsman]. Vienna: Austrian Ombudsman. Retrieved from https://strafverteidigung.at/wp-content/uploads/2019/06/Broschure_Schriftenreihe-der-Volksanwaltschaft-Band-VIII.pdf

Grabenwarter, C. (2021). StPO § 8. In H. Fuchs & E. Ratz (Eds.), *WK zur Strafprozessordnung* [Vienna Commentary on the Criminal Procedure Code]. Vienna: Manz.

Guarnera, L., Murrier, D. C., & Boccaccini, M. T. (2017). Why do forensic experts disagree? Sources of unreliability and bias in forensic psychology evaluations. *Translational Issues in Psychological Science*, *3*(2), 143-152. https://doi.org/10.1037/tps0000114

Guthrie, C., Rachlinski, J. J., & Wistrich, A. J. (2002). Judging by heuristic: cognitive illusions in judicial decision making. *Judicature*, *86*(1), 44-50.

Harris, G. T., Rice, M. E., Quinsey, V. L., & Cormier, C. A. (2015). *Violent offenders: Appraising and managing risk* (3rd ed.). Washington, DC: American Psychological Association.

Harwardt, F., & Schneider-Njepel, V. (2013). LSI-R – Level of Service Inventory-Revised. In M. Rettenberger & F. von Franqué (Eds.), *Handbuch kriminalprognostischer Verfahren* [Handbook of violence risk assessment], (pp. 243-255). Göttingen: Hogrefe.

Jerabek, R., & Ropper, R. (2024). StGB §43. In F. Höpfel & E. Ratz (Eds.), *Wiener Kommentar zum Strafgesetzbuch* [Vienna Commentary on the Criminal Procedure Code] (2nd ed.). Vienna: Manz.

Kabzinska, J. (2021). The views of judges and prosecutors on the opinions of psychologist expert witnesses in cases concerning witness testimony. *Problems of Forensics Sciences*, *125*, 67-82. https://doi.org/10.4467/12307483PFs.20.005.14785

Kaiser, N. (2023). Die bedingte Strafnachsicht nach § 43 StGB [Conditional remission of sentences according to § 43 StGB]. *Journal für Strafrecht*, *4*, 297-302.

Kaiser, N., & Leibetseder, I. (2024). Spezialprävention in der Praxis: Zum Entscheidungsverhalten von Richter:innen und Staatsanwält:innen [Specific deterrence in practice: On the decision-making behavior of judges and public prosecutors], *Journal für Strafrecht*, *2*, 125–130.

Kastner, P. (2019). Mindeststandards für forensisch-psychiatrische und psychologische Gutachten. [Minimum standards for forensic psychiatric and psychological assessments] In G. Brinek. (Ed.), *Gutachten als Schlüsselfaktoren im Maßnahmenvollzug. Schriftenreihe der Volksanwaltschaft* [Expert witness reports as key factors in the correctional system] Series of publications from the Austrian Ombudsman]. Vienna: Austrian Ombudsman. Retrieved from https://strafverteidigung.at/wp-content/uploads/2019/06/Broschure_Schriftenreihe-der-Volksanwaltschaft-Band-VIII.pdf

Kert, R. (2017). TilGG § 1. In H. Fuchs & E. Ratz (Eds.), *WK zur Strafprozessordnung* [Vienna Commentary on the Criminal Procedure Code]. Vienna: Manz.

Kovera, M. B., & McAuliff, B. D. (2000). The effects of peer review and evidence quality on judge evaluations of psychological science: Are judges effective gatekeepers? *Journal of Applied Psychology*, *85*(4), 574-586. https://doi.org/10.1037/0021-9010.85.4.574

Leibetseder, I., Kaiser, N., & Woschizka, J. (2024). Disziplinäre Grenzen und interdisziplinäre Brücken spezialpräventiver Strafzumessungspraxis [Disciplinary boundaries and interdisciplinary bridges of specific deterrence sentencing practice], *Journal für Strafrecht*, *4*(11), 341-344. https://doi.org/10.33196/jst202404034101

Müller-Isberner, R., Jöckel, D., & Gonzales Cabeza, S. (1998). *Die Vorhersage von Gewalttaten mit dem HCR-20* [The prediction of violent acts with the HCR-20.] Haina: Institut für Forensische Psychiatrie.

Müller-Isberner, R., Schmidbauer, W. M., & Born, P. (Hrsg.). (2014). *Die Vorhersage von Gewalttaten mit dem HCR-20 V3: Benutzerhandbuch Deutsche Version* [The prediction of violent acts with the HCR-20 Version 3: Manual German version] Haina: Institut für Forensische Psychiatrie.

Murrie, D. C., Boccaccini, M. T., Turner, D. B., Meeks, M., Woods, C., & Tussey, C. (2009). Rater (dis)agreement on risk assessment measures in sexually violent predator proceedings: Evidence of adversarial allegiance in forensic evaluation? *Psychology, Public Policy, and Law*, *15*(1), 19-53. https://doi.org/10.1037/a0014897

Neal, T. M. S., & Grisso, T. (2014). Assessment practices and expert judgement methods in forensic psychology and psychiatry. *Criminal Justice and Behavior*, *41*(12), 1406–1421. https://doi.org/10.1177/0093854814548449

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. (2006). *Violent offenders: Appraising and managing risk* (2nd ed.). American Psychological Association.

Rettenberger, M., Hertz, P. G., & Eher, R. (2017). Die deutsche Version des Violence Risk Appraisal Guide- Revised (VRAG-R) [The German version of the Violence Risk Appraisal Guide-Revised (VRAG-R)]. In *Elektronische Schriftenreihe der Kriminologischen Zentralstelle Band 8* [Electronic series of publications from the center of criminology. Vol. 8]. Wiesbaden: KrimZ. Retrieved from https://www.krimz.de/fileadmin/dateiablage/E-Publikationen/BM-Online/bm-online8.pdf

Rice, M. E., Harris, G. T., & Lang, C. (2013). Validation of and revision to the VRAG and SORAG: The Violence Risk Appraisal Guide – Revised (VRAG-R). *Psychological Assessment*, *25*(3), 951-965. https://doi.org/10.1037/a0032878

Riffel, R. (2023). StGB §32. In F. Höpfel & E. Ratz (Eds.), Wiener Kommentar zum Strafgesetzbuch. [Vienna Commentary on the Criminal Procedure Code]. (2nd ed.) Vienna: Manz.

Rossegger, A., Gerth, J., & Endrass, J. (2013). VRAG – Violence Risk Appraisal Guide. In M. Rettenberger & F. von Franqué (Eds.), *Handbuch kriminalprognostischer Verfahren* [Handbook of violence risk assessment], (pp. 141-158). Göttingen: Hogrefe.

Rossegger, A., Laubacher, A., Moskvitin, K., Villmar, T., Palermo, G. B., & Endrass, J. (2011). Risk assessment instruments in repeat offending: The usefulness of FOTRES. *International Journal of Offender Therapy and Comparative Criminology*, *55*(5), 716-731. https://doi.org/10.1177/0306624x09360662

Schmidt-Atzert, L., Krumm, S., & Amelang, M. (2021). Diagnostik in weiteren Anwendungsfeldern. [Diagnostics in other fields of application] In L. Schmidt-Atzert et al. (Eds.) *Psychologische Diagnostik* [Psychological Assessment], (pp. 731-781). Berlin: Springer Nature.

Schroll, H. V., & Oshidari, B. (2024). JGG § 5. In F. Höpfel & E. Ratz (Eds.), *Wiener Kommentar zum Strafgesetzbuch* [Vienna Commentary on the Criminal Procedure Code] (2nd ed.) Vienna: Manz.

Shapiro, D. L., Mixon, L., Jackson, M., & Shook, J. (2015). Psychological expert witness testimony and judicial decision making trends. *International Journal of Law and Psychiatry*, *42-43*, 149-153. https://doi.org/10.1016/j.ijlp.2015.08.020

Tadei, A., Finnilä, K., Korkman, J., Salo, B., & Santtila, P. (2014). Features used by judges to evaluate expert witnesses for psychological and psychiatric legal issues. *Nordic Psychology*, *66*(4), 239-253. https://doi.org/10.1080/19012276.2014.963648

Tadei, A., Finnilä, K., Reite, A., Antfolk, J., & Santtila, P. (2016). Judges‘ capacity to evaluate psychological and psychiatric expert testimony. *Nordic Psychology*, *68*(3), 204-217. https://doi.org/10.1080/19012276.2015.1125303

Walton, J. S. (2022). Designing rehabilitative interventions in forensic settings. In N. Gredecki & P. Turner (Eds.), *Forensic Psychology* (pp. 73-86). New York, NY: Routledge.

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20: Assessing risk for violence, version 2*. Burnaby, BC: Mental Health, Law, and Policy Institute, Simon Fraser University.

Wistrich, A. J., Rachlinski, J. J., & Guthrie, C. (2015). Heart versus head: Do judges follow the law or follow their feelings? *Texas Law Review*, *93*(4), 855–923.

AUTHORS' INFORMATION

Full name: Martina Feldhammer-Kahr, MSc, PhD
Institutional affiliation: Department of Psychology & Department of Criminal Law, Criminal Procedure Law and Criminology, University of Graz (Austria)
Institutional address: Universitätsplatz 2, 8010 Graz, Austria
Short biographical sketch: Martina Feldhammer-Kahr is currently Senior lecturer and scientist at the Department of Psychology at the University of Graz, where she earned her Master of Science and PhD. She is specialized in psychological assessment and methods and team member of Prof. Arendasy's section. The team is responsible for several projects (e.g. admission test MedAT for Human and Dental Medicine) where they use modern test construction. One focus of the research group is fairness in psychological assessments. In 2011 she participated at the European Forum Alpbach, where she worked as Assistant for Prof. Reinhard Merkel, Professor in Criminal Law and Legal Philosophy. The interdisciplinary exchange led to a special interest in Legal Psychology. In 2015 she lectured at the Okan University in Istanbul as part of an exchange program. Since 2022 she also lectures at the Department of Criminal Law, Criminal Procedure Law and Criminology of the University of Graz.

Full name: Nina Kaiser, Mag. Dr.
Institutional affiliation: Hans Gross Centre for Interdisciplinary Criminal Sciences, Department of Criminal Law, Criminal Procedure Law & Criminology, Faculty of Law, University of Graz
Institutional address: Universitätsstraße 15, B3, 8010 Graz (Austria)
Short biographical sketch: Mag. Dr. Nina Kaiser is project manager and scientist at the Hans Gross Centre for Interdisciplinary Criminal Sciences at the Institute of Criminal Law, Criminal Procedure and Criminology at the Faculty of Law of the University of Graz. Background: Diploma and Doctoral Studies at the University of Graz Law School, 2016-2020 University Assistant at the Institute for Criminal Law, Criminal Procedure and Criminology, since February 2020 doctorate in law; since 2023 voluntary probation officer. Inspired by Hans Gross, she is primarily concerned with interdisciplinary interfaces between Criminal Law and the fields of Criminology, Criminalistics and Forensic Sciences.

Full name: Ida Leibetseder, BSc., Mag
Institutional affiliation: Hans Gross Centre for Interdisciplinary Criminal Sciences, Department of Criminal Law, Criminal Procedure Law & Criminology, Faculty of Law, University of Graz
Institutional address: Universitätsstraße 15, B3, 8010 Graz (Austria)
Short biographical sketch: Mag. Ida Leibetseder, BSc. is Praedoc and project assistant at the Hans Gross Centre for Interdisciplinary Criminal Sciences at the Institute of Criminal Law, Criminal Procedure and Criminology at the Faculty of Law of the University of Graz. Her field of research is the interface between psychology, criminology and law.

Full name: Markus Sommer, MSc, PhD
Institutional affiliation: Department of Psychology, University of Graz (Austria)
Institutional address: Universitätsplatz 2, 8010 Graz, Austria
Short biographical sketch: Markus Sommer works as a senior scientist at the section of Psychological Assessment and Methods headed by Prof. Arendasy at the Department of Psychology at the University of Graz. He earned his master's degree at the University of Vienna, and completed his PhD at the University of Graz. His research interest focus on threats to the validity and fairness of assessments, and how these threats can be circumvented using theory-driven approaches to automatic item generation. He is also part of the research team headed by Prof. Arendasy, which is responsible for several large-scale and high-stakes assessment projects (e.g. Austrian medical school admission test; vocational counselling of the chamber of commerce in Styria etc.) in Austria and abroad.

Full name: Martin Arendasy, MSc, PhD
Institutional affiliation: Department of Psychology, University of Graz (Austria)
Institutional address: Universitätsplatz 2, 8010 Graz, Austria
Short biographical sketch: Martin E. Arendasy, MSc, PhD, serves as Full Professor and Chair of the Psychological Assessment and Quantitative Methods section at the Department of Psychology, University of Graz, Austria. His principal research interests encompass the development of innovative methodologies for item and test construction, notably automated item generation. He has been awarded a patent in the domain of 'two-component automatic item generation' and has received several research accolades, including recognition by the German Association of Research Psychologists.
Prof. Arendasy's applied work primarily focuses on high-stakes large-scale assessments, such as the medical college admission test (MedAT) for Austrian Medical Universities, and the development of advanced test batteries used in Talent Centers for adolescents aged 13-15, which are operated by the chambers of commerce in Austria and Italy. His future research directions include securing a second patent in automated item and test construction and investigating the responsible and methodologically rigorous application of deep learning techniques across diverse academic and practical contexts.