



AN EVALUATION OF WITNESS TESTIMONY IN CRIMINAL PROCEEDINGS: PRELIMINARY RESULTS OF EMPIRICAL RESEARCH

Joanna KABZIŃSKA

*SWPS University of Social Sciences and Humanities, Warsaw, Poland
Katowice Faculty of Psychology, Katowice, Poland*

Abstract

The research paper investigates the assessment of the legal credibility of eyewitness testimony in criminal proceedings. It examines the criteria used by courts to assess eyewitness credibility and the factors that influence the number and type of criteria they invoke in their evaluations. It also raises the question of whether the assessment of credibility is subject to bias. The research aims were achieved by analysing the testimony of 751 eyewitnesses examined in criminal trials (N = 90). The assessment of legal credibility was found to be superficial: the courts invoke a low overall number of criteria to justify their judgments and, despite individual differences in testimony and eyewitnesses, tend to refer to only a small number of criteria. Numerous factors, which are connected with both the procedure and the eyewitnesses themselves, influence the legal assessment of eyewitness reliability. The analysis demonstrated that the assessment of legal credibility is subject to certain biases. The paper concludes by discussing the practical implications of the results and suggesting avenues for further research.

Keywords

Credibility of eyewitness evidence; Legal credibility assessment; Eyewitness testimony; Judicial decision-making.

Received 6 August 2018; accepted 10 September 2018

Introduction

To establish the facts of court proceedings it is necessary to assess all of the evidence gathered and decide what connects it. The principle that evidence may be freely assessed, which is expressed in article 7 of the Code of Criminal Procedure, upholds this model of adjudication and compels the courts to form an opinion based on all the evidence presented aided by correct reasoning, knowledge and life experience. To evaluate evidence in accordance with scientific knowledge means to do so in accordance with reliable, up-to-date research findings. In particular this means following the scientific method and taking advantage of its advances (Waltoś, Hofmański, 2013). Knowledge of the psychology of witness testimony is especially

useful in the free assessment of the evidence derived from it. Though logic is without doubt the basic evidence-assessment tool at the courts' disposal, a knowledge of psychology is, for a lawyer, an essential tool of the trade that guarantees the impartiality and accuracy of decisions (Gruza, 2014). The criteria for evidence assessment have been studied by a number of scholars of psychology and criminology, including Gruza (2003), Maciejewski (2009) and Płachta (1985). Yet knowledge of evidence assessment is not widespread among practitioners and is in many cases non-existent (Gruza, 2014). It is difficult to acknowledge the examples adduced in case law of which criteria to be guided by (and more often of which criteria not to be guided by) as exhaustive. See, for example, judgment SN (Supreme Court) of 11 January 1996, II KRN 178/95.

Comparatively little is known about the influence of the psychology of witness statements on judicial practice. The knowledge participants in criminal proceedings have of the psychology of witness statements has hitherto been studied in surveys. These American (Wise, Safer, 2004), Norwegian (Magnussen et al., 2008) and Swedish (Granhag, Strömwall, Hartwig, 2005) investigations of judges found that their subjects knew little of the factors and processes influencing the perception and memory of eye witnesses. Yet research has found that the psychological knowledge judges have has a positive influence on their rulings (Wise, Safer, 2004). What is more, the education of judges and other participants in criminal proceedings has been described in more than one study (Wise, Dauphinais, Safer 2007, for example) as a significant factor in minimising untrue or false statements in criminal trials. There have been very few similar research endeavours aimed at establishing the psychological knowledge of professional Polish participants in criminal proceedings and its influence on them (see Piekarska-Drażek, 2004). Yet simply determining that lawyers possess knowledge of a specific field does not mean that they have the capacity to exercise it skilfully.

Sporer (2008) found that previous studies of the psychology of witness statements had almost entirely overlooked the question of how the testimonies are evaluated. He therefore recommended that researchers focus on the way in which the judiciary assesses witness statements (Sporer, 2008). To this end, the present study examines the practice of evaluating the reliability of witness testimony by analysing the records of criminal cases. The following research questions were posed:

In practice, what criteria do judges use to evaluate witness statements?

What factors influence the number and type of criteria applied to determine the reliability of testimony?

What factors influence the final result of an assessment of witness testimony?

Materials and methods

Criminal cases from two courts within the local jurisdiction of the Kraków Appeal Court – Brzesko District Court and the District Court for Kraków-Nowa Huta – were examined¹. A request was made to both courts to

¹ The research will ultimately be extended to embrace eight district courts in Poland from within the local jurisdictions of the appeal courts in Kraków, Poznań, Lublin and Gdańsk.

make available the records of a total of fifty-four cases that meet the following criteria:

- a) cases completed with legal validity in 2013–2015,
- b) cases in which at least one witness was examined,
- c) cases in which the court of first instance had written a reasons for judgement,
- d) cases of low-severity, medium-severity and high-severity crime².

Crimes were selected for these categories according to which of the twenty-two chapters of the special section of the penal code satisfied the following conditions:

- a) the greatest number of convictions in 2015 (but no fewer than 1,500 people),
- b) cases that in the main relied on evidence other than witness testimony were eliminated.

Seven chapters of the penal code were selected as a result of applying these conditions. A maximum of three crimes that produced the greatest number of convictions in 2015 (but no fewer than 500 people) was selected from each chapter. Fifteen crimes were selected as a result of this analysis. Of these, three were discarded during pilot studies. Four crimes were selected for each level of severity. High severity: a) sexual activity with a minor under fifteen years old, b) active bribery, c) fraud, d) robbery. Medium severity: a) actual bodily harm, b) coercion to force repayment of a debt, c) cruelty, abuse and bullying, d) theft. Low severity: a) minor health impairment, b) punishable threat, c) assaulting a public servant, d) defamation and insult of a public servant.

Paying particular attention to the courts' reasons for judgments, the case records were analysed using a three-part questionnaire concerning the case, the witness and the evaluation of the testimony. The reliability-assessment criteria were selected based on the available literature (Gruza, 2003; Maciejewski, 2009; Płachta, 1985), judgments of the Polish appeal courts and supreme court, and the pilot studies for the questionnaire (N = 28). They were divided into objective criteria relating to the content of the statements and subjective criteria relating to the person of the witness. Table 1 contains information on the legal criteria for the evaluation of testimony in the questionnaire with an explanation for each criterion.

² The term of imprisonment to be served for committing it – as set out in the Penal Code – was taken as the measure of a crime's degree of seriousness. A low-severity crime is one that attracts a sentence of up to three years' imprisonment. A medium-severity crime is one that attracts a sentence of at least three, but no more than five, years' imprisonment. A high-severity crime is one that attracts a sentence of more than five years.

Table 1
Objective and subjective criteria defined in the research. The non-credibility criteria are presented in italics

General criterion	Credible and non-credible version of the criterion	Explanation of the criterion
Objective criteria		
Inner consistency	Internal consistency of testimony	Eyewitness testimony is consistent, consequent, constant
	<i>Internal inconsistency of testimony</i>	Eyewitness testimony is inconsistent, inconsequent and contains discrepancies
Logic	Logical testimony	Eyewitness testimony is logical, understandable, comprehensible, makes sense and is clear
	<i>Lack of logic in the testimony</i>	Eyewitness testimony is illogical, incomprehensible, makes no sense and lacks clarity
External coherence	Coherence with other evidence	Eyewitness testimony is coherent with other evidence presented in court
	<i>Incoherence with other evidence</i>	Eyewitness testimony is inconsistent or conflicts with other evidence presented in the case
Number of details	Detailed testimony	Eyewitness testimony is detailed, exact; the amount of details provided by the eyewitness is adequate
	<i>Insufficient number of details</i>	Eyewitness testimony lacks detail; it is vague and inaccurate; the number of details is inadequate
Substantiality	Substantial testimony	Eyewitness testimony is concrete, substantial, tangible, concentrated on the merits of the case
	<i>Insubstantial testimony</i>	Eyewitness testimony is non-specific, insubstantial, lacks concentration on the merits of the case
Life experience	Corresponding with life experience	Eyewitness testimony is rational, consistent with life experience
	<i>Not corresponding with life experience</i>	Eyewitness testimony is irrational and contradicts life experience
Objectivism	Objectivism in the testimony	The eyewitness provides information that is unfavourable to him/her (or people familiar to him/her); admits to a justifiable lack of memory of the event
	<i>Lack of objectivism in the testimony</i>	The eyewitness conceals information that is unfavourable to him/her (or people familiar to him/her); the eyewitness makes himself/herself look good
Context	Contextual embedding	Eyewitness testimony is embedded in place and time, in the eyewitnesses' daily routine; the testimony is embedded in a personal and/or emotional context
	<i>Lack of contextual embedding</i>	Eyewitness testimony lacks embedding in the spatial, temporal, personal or emotional context
Source of the testimony	Direct account	Eyewitness testimony refers to events which he/she directly witnessed or was involved in personally
	<i>Indirect account</i>	Eyewitness testimony refers to events which he/she did not witness directly
Lack of undermining premises	No premises to undermine the credibility of the testimony	The testimony was not discredited
Visibility	Objective possibility to make the observation	Eyewitness had an objective opportunity to observe the event
	<i>No objective possibility to make the observation</i>	Eyewitness had no objective opportunity to observe the event
Other	Other objective credibility criterion	Any other objective credibility criterion not listed above
	<i>Other objective non-credibility criterion</i>	Any other objective non-credibility criterion not listed above

General criterion	Credible and non-credible version of the criterion	Explanation of the criterion
Subjective criteria		
Spontaneity	Unstructured/Spontaneous manner of testifying	Eyewitness testifies spontaneously, freely
	<i>Structured (prepared in advance) manner of testifying</i>	Eyewitness testimony lacks spontaneity, the testimony gives the impression of having been prepared in advance
Confidence	Confidence of the eyewitness	Eyewitness is confident, categorical
	<i>Lack of confidence in the eyewitness</i>	Eyewitness lacks confidence, gives the impression that his/her testimony is wavering
Motivation arising from familiarity with the defendant	Lack of relationship with the defendant	Lack of relationship between the eyewitness and the defendant on any ground: personal, emotional or professional that could offer evidence of his/her credibility
	<i>Relationship with the defendant</i>	Presence of a personal, emotional or professional relationship between the eyewitness and the defendant which suggests lack of eyewitnesses' credibility
Motivation arising from interest in the outcome	Lack of interest in the outcome of the trial	Eyewitness is not interested in a particular outcome of the court case (sentencing or acquittal)
	<i>Interest in a particular outcome of the trial</i>	Eyewitness testifies in a manner that is supposed to lead to a particular outcome of the case (e.g. in the testimony is pursuing a defence strategy arranged by the defendant)
Other types of motivation	Other – credible – motivation to testify	Lack of reasons to testify falsely or untruthfully. The reasons behind the decision to reveal the crime/to testify/to provide the court with a particular testimony are justifiable
	<i>Other non-credible motivation to testify</i>	There are reasons to suggest that the eyewitness testified falsely or untruthfully. The reasons behind the decision to reveal the crime/to testify/to provide the court with particular testimony raise doubts
Profession	Eyewitnesses' profession	The profession of the eyewitness suggests his/her credibility (e.g. police officer)
Other	Other subjective credibility criterion	Any other subjective credibility criterion not listed above
	<i>Other subjective non-credibility criterion</i>	Any other subjective non-credibility criterion not listed above

Sample

Ninety criminal cases from two courts within the local jurisdiction of the Kraków Appellate Court – Brzesko District Court (N = 44) and the District Court for Kraków-Nowa Huta (N = 46) – were examined. Table 2 contains detailed information on the structure of the sample.

A total of 751 witnesses were questioned in the analysis (at Brzesko District Court: N = 361; at the District Court for Kraków-Nowa Huta: N = 390). The sample contained 292 women (38.9%) and 459 men (61.1%). Only two witnesses aged fifteen or younger were questioned, while seventy witnesses aged over sixty-five were questioned. Table 3 contains detailed information on the structure of the sample of witnesses.

Results

General results of the courts' evaluation of testimony

The average number of witnesses questioned in the cases analysed was 8.34 (SD = 6.15). No assessment was made by either of the courts of the evidence of 104 of the overall total of 751 witnesses. The reasons given for refusing to make an assessment were varied. The predominant justification was that the witness possessed no information relevant to the case. In some instances the courts did not raise at all the question of the reasons for not making an assessment³. The witness testimonies were found to be credible in the majority of cases (Brzesko District Court: N = 229;

³ The 104 cases for which no assessment of testimony was made were excluded from further analysis.

Table 2

The number and percentage of analysed criminal cases according to the severity of the offence, the court's final judgment and the severity of the final judgment

Inferior court	Cases according to the severity of the offence			Cases according to the final judgment		Cases according to the severity of the final judgment**		
	Low severity offence	Medium severity offence	High severity offence	Conviction	Acquittal	Judgment less severe than requested by the public prosecution	Judgment identical to the request of the public prosecution	Judgment more severe than requested by the public prosecution
Brzesko (small court*)	15	15	14	35	9	30	4	3
Kraków-Nowa Huta (large court*)	10	15	21	45	1	26	12	6

* – A small court (in this instance Brzesko) is an inferior court which is located in a city which is not simultaneously the location of a provincial court. A large court is an inferior court located in a city which is the location of a provincial court as well (in this instance Kraków).

** – The severity of the final judgment was established by comparing the request of the public prosecution (expressed at the final trial hearing before the delivery of the final verdict) with the final judgment. The overall number of cases according to the severity of the penalty is lower than in other instances due to the fact that in several cases the public prosecutor did not express a request concerning the final judgment.

Table 3

The number and percentage of eyewitnesses interviewed at trial in the analysed cases according to their procedural characteristics

	Role in the event (and the proceedings)			Relationship with the defendant			Initiative to appoint the witness		Expert psychologist present at the interview	
	Victim	Bystander	No data available	Unrelated	Related*	No data available	Defendant	Other	Present	Not present
Small court	62	298	1	299	61	1	33	328	8	353
Large court	90	300	0	331	54	5	35	355	3	387

* – Being related to the defendant was assessed based on whether the eyewitness was entitled by the Polish C.C.P. to refuse to testify in the case of a particular defendant – i.e. when the eyewitness is the defendant's spouse or life partner, parent (or grandparent etc.), child (or grandchild etc.), sibling, parent in law, as well as brother- (or sister-) in-law (spouses and in-laws are entitled to refuse to testify despite prior dissolution of marriage).

Table 4

The number and percentage of objective credibility-criteria invoked by the court. Non-credibility criteria are presented in italics

		Objective criteria indicated as the grounds of the assessment	
		Inferior court Brzesko	Inferior court Kraków-Nowa Huta
		<i>N</i>	<i>N</i>
Inner consistency	Internal consistency of testimony	54	138
	<i>Internal inconsistency of testimony</i>	27	34
Logic	Logical testimony	30	146
	<i>Lack of logic in the testimony</i>	6	14
External coherence	Coherence with other evidence	198	237
	<i>Incoherence with other evidence</i>	34	32

		Objective criteria indicated as the grounds of the assessment	
		Inferior court Brzesko	Inferior court Kraków-Nowa Huta
		<i>N</i>	<i>N</i>
Number of details	Detailed testimony	33	35
	<i>Insufficient number of details</i>	18	12
Substantiality	Substantial testimony	35	63
	<i>Insubstantial testimony</i>	8	3
Life experience	Corresponding with life experience	6	13
	<i>Not corresponding with life experience</i>	4	7
Objectivism	Objectivism in the testimony	75	49
	<i>Lack of objectivism in the testimony</i>	13	9
Context	Contextual embedding	2	4
	<i>Lack of contextual embedding</i>	0	0
Source of the testimony	Direct account	8	22
	<i>Indirect account</i>	14	11
Lack of undermining premises	No premises to undermine the credibility	31	26
Visibility	Objective possibility to make the observation	3	1
	<i>No objective possibility to make the observation</i>	7	4
Other	Other credibility criterion	3	1
	<i>Other non-credibility criterion</i>	1	1
Overall		610	862

District Court for Kraków-Nowa Huta: $N = 284$). But in 134 cases they were found to be unreliable or partially unreliable (Brzesko District Court: $N = 70$; District Court for Kraków-Nowa Huta: $N = 64$). In the cases analysed, the courts drew upon 1,723 criteria for assessing witness testimony in their reasons for judgment. The average number of criteria drawn upon to evaluate a witness statement was 2.65 ($SD = 1.7$). The evaluations were dominated by objective criteria referring to the content of the testimonies. A total of 1,472 objective criteria were applied ($M = 2.26$; $SD = 1.49$) and 251 subjective criteria relating to the person of the witness ($M = 0.39$; $SD = 2.26$). Not all of the testimonies analysed referred to a specific type of criteria. Objective criteria were mentioned in 599 of the testimonies evaluated ($M = 2.46$; $SD = 1.39$), while subjective criteria were mentioned in as few as 196 ($M = 1.28$; $SD = 0.55$). Table 4 and Table 5 show the frequency with which objective and subjective criteria are mentioned for both courts in the analysis.

For testimony deemed credible at both courts (analysed together), the criterion of consistency of testimony with other evidence in the case ($N = 435$) appeared most frequently in reasons for judgment. The second and third most frequent criteria were internal

consistency of evidence ($N = 192$) and the logicity of testimony ($N = 176$). This was the exact triumvirate of the most frequently applied objective criteria at the District Court for Kraków-Nowa Huta (but with second and third places reversed). The criterion the judges at Brzesko District Court mentioned most often was consistency of testimony with other evidence in the case, which was followed by internal consistency of evidence and, in third place, the objectivity of testimony. Where witness statements were not regarded as credible, or were regarded as partially credible, the most frequently mentioned criteria for evaluating this lack of credibility were inconsistency with the evidence, inconsistent testimony and over-generalised testimony.

For testimony deemed credible at both courts (analysed together), the subjective criteria mentioned most often in justifications of rulings were other credible motivation to testify ($N = 66$), followed by lack of relationship with the defendant ($N = 43$) and unstructured manner of testifying ($N = 11$). There were, however, differences between the two courts. While the first two motivational criteria mentioned above were the ones most frequently mentioned in the justifications of both courts, the third most frequently men-

Table 5

The number and percentage of subjective credibility criteria invoked by the court. Non-credibility criteria are presented in italics.

		Subjective criteria indicated as the grounds of the assessment	
		Inferior court in Brzesko	Inferior court in Kraków-Nowa Huta
		<i>N</i>	<i>N</i>
Spontaneity	Unstructured manner of testifying	9	2
	<i>Structured manner of testifying</i>	0	0
Confidence	Confidence of the eyewitness	0	4
	<i>Lack of confidence of the eyewitness</i>	0	0
Motivation arising from familiarity with the defendant	Lack of relationship with the defendant	31	12
	<i>Relationship with the defendant</i>	27	22
Motivation arising from interest in the outcome	Lack of interest in the outcome of the trial	8	2
	<i>Interest in particular outcome of the trial</i>	6	1
Other types of motivation	Other – credible – motivation to testify	49	17
	<i>Other non-credible motivation to testify</i>	15	24
Profession	Eyewitnesses' profession (e.g. police officer)	2	4
Other	Other subjective credibility criterion	1	6
	<i>Other subjective non-credibility criterion</i>	6	3
Overall		154	97

tioned criteria were different. At the District Court for Kraków-Nowa Huta it was other criterion of credibility and at Brzesko District Court it was unstructured manner of testifying.

Selected independent variables and the number of criteria of testimony evaluation mentioned in the reasons for judgment

Analysis of the results indicates a significant statistical relationship between certain attributes of proceedings and characteristics of witnesses and the number of criteria for evaluating the credibility of testimony judges mention in reasons for judgment (Table 6).

A statistical analysis using the non-parametric Kruskal-Wallis Test indicates a statistically significant difference between the number of criteria mentioned and the seriousness of the crime ($H = 15.32$, $p < 0.05$). Comparisons were made using the Mann-Whitney U Test to find which of the differences were statistically significant. Judges employ fewer criteria of credibility to analyse testimony in high severity cases when compared with medium severity ($Z = -2.57$, $p < 0.05$) and low severity ($Z = -3.92$, $p < 0.05$) cases. A statistically significant relationship between the number of testimony-evaluation criteria mentioned and the

severity of the penalty adjudicated by the court was also observed ($H = 19.26$, $p < 0.05$). Further analysis using the Mann-Whitney U Test indicated that judges refer to significantly fewer criteria when the penalty is more severe than that requested by the public prosecutor than they do when the penalty is either the same as that requested by the public prosecutor ($Z = -4.315$, $p < 0.05$) or less severe than that requested by the public prosecutor ($Z = -3.78$, $p < 0.05$).

The results indicated a statistically significant relationship between all of the following bracketed material (the role of the witness in the proceedings, the age of the witness, the relationship between the witness and the accused) and the number of evaluation criteria mentioned in the reasons for judgment. Judges mention a greater number of criteria in assessments of the testimony of victims than they do in assessments of witness testimony ($Z = -7.11$, $p < 0.05$). Furthermore, an r-Pearson Correlation indicated a weak but statistically significant negative relationship between the age of the witness and the number of criteria mentioned ($r = -0.1$, $p < 0.05$). What is more, the results of the analyses suggested a statistically significant difference in the number of credibility-of-testimony evaluation criteria mentioned by the court depending on the relationship between the witness and the accused. Judges

Table 6

Mean number and standard deviation (in brackets) of credibility criteria invoked by the courts in justifications of rulings and the values for statistical tests (Kruskall-Wallis – in italics – or Mann-Whitney U test)

Variables		M (SD)	<i>H/Z</i>
Severity of the offence	Low severity	2.98 (1.76)	<i>15.32**</i>
	Medium severity	2.72 (1.81)	
	High severity	2.25 (1.39)	
Severity of the final judgment	Less severe	2.67 (1.67)	<i>19.26**</i>
	Identical to	2.94 (1.89)	
	More severe	1.80 (1.27)	
Degree of involvement in the event	Bystander	2.36 (1.51)	<i>-7.11**</i>
	Victim	3.6 (1.95)	
Relationship with the defendant	Related	3.35 (1.92)	<i>-4.32**</i>
	Unrelated	2.51 (1.63)	
Expert psychologist presence at interview	Present	4.22 (0.83)	<i>-3.24*</i>
	Not present	2.62 (1.71)	
Final assessment of the eyewitnesses' credibility	Credible	2.50 (1.68)	<i>-4.96**</i>
	Not credible	3.26 (1.67)	

* – p value < 0.001

** – p value < 0.0001

use a greater number of criteria of credibility to assess the statements of people related to the accused than they do to assess those who are unrelated to, and do not know, the accused ($Z = -4.32$, $p < 0.05$). The participation of a psychologist in witness examinations also has an influence on the number of testimony-assessment criteria mentioned in reasons for judgment. Judges use criteria for testimony evaluation significantly more often to assess the testimony of witnesses in proceedings involving a psychologist (who also writes a report on the psychological validity of their testimony) than where a psychologist is not present during examination of witnesses ($Z = -3.24$, $p < 0.05$). Finally, a statistically significant difference was found in the number of reliability-assessment criteria applied depending on the court's final assessment of the testimony. Judges mention a greater number of criteria in reasons for judgment when they deem the testimony to be unreliable (or partially unreliable) than when they deem it reliable ($Z = -4.96$, $p < 0.05$).

Selected independent variables and the type of testimony-evaluation criteria mentioned in reasons for judgment.

The types of criteria mentioned by judges in reasons for judgment (objective versus subjective) were also investigated. First of all, a comparison was made to determine whether selected dependent variables influence the relationship of the subjective to the objective criteria mentioned in evaluations of testimony (Table 7).

The analysis found that the size of the court (small versus large) in which the ruling is made influences the choice of criteria made by judges when justifying their assessments of testimony⁴. Evaluations of testimony justified by reference to more subjective criteria than objective criteria appeared more often (at a statistically significant level) in small courts (6.6%) than in large courts (1.1%). At the same time, evaluations of testimony that referred to an equal number of subjective and objective criteria were made more often (at

⁴ An analysis was carried out in two courts: a large district court (District Court for Kraków-Nowa Huta) understood as a court located in a place in which a provincial court also sits, and a small district court (Brzesko District Court), which is located in a place in which a provincial court does not sit.

Table 7

The percentage proportions of the two types of credibility criteria invoked by the courts to assess the credibility of eyewitness testimony

		Proportion of objective and subjective credibility criteria			
		No criteria	Objective criteria predominate	Subjective criteria predominate	Equal number of both types of criteria
Size of the court	Small court	4.6	76.9	6.6*	11.9*
	Big court	6.0	89.4	1.1*	3.4*
Severity of the offence	Low severity	2.9	90.1*	3.5	5.1*
	Medium severity	6.7	78.7*	3.2	11.3*
	High severity	5.6	84.8*	4.6	3.5*
Degree of involvement in the event	Bystander	6.6	79.9*	4.6	8.9
	Victim	1.4	95.9*	0.7	2.0
Relationship with the defendant	Unrelated	6.3*	83.6	3.9	6.3*
	Related	0.9*	83.0	2.8	13.2*
Final assessment of the eyewitness' credibility	Credible	6.4*	86.0*	3.5	4.1*
	Not credible	0.7*	76.1*	3.7	19.4*

* – p value < .05

a statistically significant level) in small courts (11.9%) than in large courts (3.4%, $\chi^2(3) = 32.29$, $p < 0.05$).

In the case of low-severity crimes, testimony was evaluated significantly more often by reference to a majority of objective criteria over subjective criteria (90.1%) when compared to testimony lodged in the case of medium-severity (78.7%) and low-severity crimes (84.8%). In connection with medium-severity crimes, however, judges justified their evaluations twice as often by reference to the same number of objective criteria as subjective criteria than in connection with the other two categories of severity of crimes ($\chi^2(6) = 16.24$, $p < 0.05$).

It was also found that evaluations of testimony given by victims were justified significantly more often by reference to a majority of objective criteria over subjective criteria (95.9%) when compared to testimony lodged by disinterested witnesses (79.9%, $\chi^2(3) = 21.21$, $p < 0.05$). Assessments of the testimony of witnesses that had a family connection with the accused in which the court did not refer to any criteria were significantly rarer (0.9%) when compared to assessments of witnesses unrelated to the accused (6.3%). In addition, the testimony of witnesses related to the accused was evaluated significantly more often using an equal number of objective and subjective criteria (13.2%) when compared to the testimony of witnesses unrelated to the accused (6.3%, $\chi^2(3) = 10.69$, $p < 0.05$).

The final assessment of witness statements as reliable or unreliable is a further factor that has an impact on the proportion of subjective to objective criteria in evaluations of testimony. Evaluations of testimony that did not refer to any criteria were made significantly more often in the case of testimony deemed reliable (6.4%) than in the case of testimony deemed unreliable (0.7%). Assessments of testimony regarded as reliable were justified significantly more often by reference to a majority of objective criteria over subjective criteria when compared to testimony regarded as unreliable. The evaluation of unreliable testimony was justified significantly more often by reference to the same number of objective as subjective criteria (19.4%) when compared with reliable testimony (4.1%, $\chi^2(3) = 42.06$, $p < 0.05$). An in-depth analysis of the average number of subjective and objective criteria mentioned in reasons for judgment was then performed.

The analysis found statistically significant differences in the average number of objective and subjective criteria employed depending on the size of the court making the assessment ($Z = -4.19$, $p < 0.05$ and $Z = -5.03$, $p < 0.05$ respectively). In other words, judges sitting in a large court (in a district court in a place that is also the seat of a provincial court) employ a greater number of objective criteria and fewer subjective criteria than judges sitting in a small court (in a place that is not also the seat of a provincial court). Furthermore, statistically significant differences were identified in the average number of objective and sub-

Table 8

Mean number and standard deviation (in brackets) of objective and subjective credibility criteria invoked by the courts in reasons for judgment and the values of statistical tests (Kruskall-Wallis – in italics – or Mann-Whitney U test)

Variables		Objective criteria		Subjective criteria	
		<i>M (SD)</i>	<i>H (Z)</i>	<i>M (SD)</i>	<i>H (Z)</i>
Size of the court	Minor court	2.01 (1.41)	-4.19***	0.51 (0.71)	-5.03***
	Large court	2.48 (1.52)		0.28 (0.58)	
Severity of the offence	Low severity	2.57 (1.52)	<i>14.34**</i>	0.41 (0.65)	<i>11.98**</i>
	Medium severity	2.29 (1.59)		0.43 (0.64)	
	High severity	1.95 (1.24)		0.30 (0.69)	
Degree of involvement in the event	Bystander	2.00 (1.35)	-7.87***	0.37 (0.65)	-1.27
	Victim	3.16 (1.60)		0.44 (0.69)	
Relationship with the defendant	Related	2.68 (1.61)	-6.11*	0.67 (0.67)	-2.88**
	Unrelated	2.18 (1.45)		0.33 (0.64)	
Final assessment of the eyewitness' credibility	Credible	2.19 (1.45)	-2.51*	0.31 (0.64)	-7.46***
	Not credible	2.60 (1.59)		0.66 (0.62)	

* – p value < 0.05

** – p value < 0.001

*** – p value < 0.0001

jective criteria mentioned depending on the severity of the crime that was the subject of the proceedings ($H = 14.34$, $p < 0.05$ and $H = 11.98$, $p < 0.05$ respectively). The courts mention significantly fewer objective criteria in evaluations of testimony lodged in high-severity cases than in low severity-cases ($Z = -3.88$, $p < 0.05$). The difference in the average number of objective criteria mentioned in evaluations of testimony in high-severity and medium-severity cases was close to statistical significance ($Z = -1.91$, $p = 0.56$). The courts also mention significantly fewer objective criteria in evaluations of testimony lodged in medium-severity cases than in low-severity cases ($Z = -2.12$, $p < 0.05$). It was further found that the courts mention significantly fewer subjective criteria in high-severity cases than in medium-severity and low-severity cases ($Z = -3.88$, $p < 0.05$ and $Z = -2.61$, $p < 0.05$ respectively). The difference in the number of subjective criteria applied in medium-severity and low-severity cases was not statistically significant.

Significant differences in the average number of objective and subjective criteria depending on the characteristics of the witnesses examined were also found. Judges mention a significantly greater number of objective criteria in evaluations of testimony lodged by victims when compared to testimony lodged by disinterested witnesses ($Z = -7.87$, $p < 0.05$). Furthermore, the results suggest that on average judges mention

a greater number of objective criteria in evaluations of testimony lodged by witnesses related to the accused than in evaluations of testimony lodged by witnesses unrelated to the accused ($Z = -6.11$, $p < 0.05$). A similar tendency was detected with respect to subjective criteria ($Z = -2.88$, $p < 0.05$). Finally, it was noted that courts mention a significantly greater number of both objective and subjective criteria when they regard testimony as unreliable (or partially unreliable) than when they regard it as reliable ($Z = -2.51$, $p < 0.05$ and $Z = -7.46$, $p < 0.05$).

Factors influencing decisions on the reliability or unreliability of witness testimony

The search for answers to the second research question, which concerned the influence of selected factors on the final evaluation of witness testimony, revealed some interesting results. Judges regard the testimony of victims as unreliable more rarely (15.1%) than that of disinterested witnesses to a crime (22.4%, $\chi^2(1) = 3.69$, $p < 0.05$).

Moreover, all of the testimonies of minors ($N = 2$) were regarded as unreliable (or partially unreliable), while 21.1% of the testimonies of those aged over sixty-five were regarded as unreliable and 20.7% of the testimonies of adults (16–64) were regarded as unreliable. This relationship was statistically significant

($\chi^2(2) = 7.58, p < 0.05$). Yet the result should be treated with caution because the assumption of the χ^2 test was contravened (33.3% of the cells in the cross-table had an expected value below 5).

The testimony of witnesses related to the accused was regarded as unreliable significantly more often (37.7%) than the testimony of witnesses not related to the accused (17.4%, $\chi^2(1) = 22.17, p < 0.05$). A similar tendency was noted in assessments of the testimony of witnesses summoned on the initiative of the accused or of a representative of the accused. Of these testimonies, 30.6% were regarded as unreliable when compared with the testimonies of witnesses summoned on the initiative of other institutions or entities (19.9%). The difference did not, however, prove to be statistically significant.

Discussion

The average number of witnesses examined in both district courts indicates that witness testimony remains one of the most important evidential resources in criminal cases. This would seem to be confirmed by the view prevailing in the literature that witness testimony is a valuable form of evidence (see de Poot, Bokhorst, van Koppen, Muller, 2004; Waltoś, 1985, for example). Research has demonstrated that judges and juries (in the common law system) rely on evidence from witness testimony (see Semmler, Brewer, Douglas, 2011; Vallas, 2011, for example). It was beyond the scope of the current research to establish the structure of the evidence in the cases investigated. There is therefore no way to decide whether witness testimony predominates with regard to other forms of evidence, or whether greater attention is paid to it in evaluations of evidence than to other forms of evidence.

It is striking that the courts examine such a large number of witnesses and at the same time evaluate their testimony in a way that appears superficial. Firstly, more than 10% of the testimonies lodged in court sessions are not assessed by the court at all in their reasons for judgment. Secondly, the average number of testimony-assessment criteria mentioned in reasons for judgment is not imposing. In their reasons for judgment, judges employed comparatively few objective criteria – and still fewer subjective ones – and were particularly reluctant to draw upon subjective criteria. This is quite surprising in the light of the importance for the assessment of testimony of taking account both of the content of the statement and the person of the witness testifying, which is a position that has been expressed many times in the case law of the Polish courts

(judgment SN (Supreme Court) of 24 November 1975, II KR 254/75; judgment SN (Supreme Court) of 10 August 2005, II KK 410/04). Why, then, are courts unwilling to draw upon subjective criteria in assessing witness testimony? Could it be that judges are fearful of supporting their evaluations with arguments that might appear intangible or fanciful? It is comparatively easy to demonstrate that a witness statement is internally consistent (or inconsistent) with the remaining evidential material by referring to the relevant sections of the examination protocol or to other evidence. It is more difficult, however, to show that a witness was (or was not) confident or that a witness testified spontaneously (or gave the impression of having learned their lines). Despite this, judges should not abandon references to subjective criteria in their evaluations of testimony. If judges do not overlook subjective aspects such as the question of motivation to testify or the use of psychological tools to assess the psychological veracity of testimony, such as Statement Validity Analysis (Memon, Vrij, Bull, 2003), they should not overlook subjective criteria in the legal assessment of reliability either. It is clear, however, that subjective criteria should supplement assessments of the content of testimony that are based on objective criteria.

The frequency with which objective criteria are drawn upon in assessments of testimony is what ultimately demonstrates their superficiality. The objective criteria of assessment that dominated the reasons for judgment were consistency with other evidence, and the logicity and internal consistency of the testimony, while the most prominent subjective criteria were those associated with the witness's motivation to testify. There is no doubt that the courts attach the greatest importance to the criteria of consistency with other evidence. They are probably encouraged in this by the position of the Supreme Court, which stated in one of its rulings that the prime factor in deciding the reliability of witness testimony is whether it is confirmed in the other evidence at the disposal of the court. Other circumstances only become important in the assessment of testimony when such confirmation is absent (judgment of the Supreme Court of 28 May 1979, I KR 91/79). It would appear, however, that this is an over-simplified position that takes account neither of the complexity of the witness statements themselves or of the process of assessing them. Such a belief, if indeed the judges do share it, could contribute in some measure to a lack of regard for the importance of assessing testimony. If it is only confirming its inconsistency with the other evidence that triggers a more penetrating assessment of testimony, then the assessments will remain superficial and generalised for as long as

this obvious contradiction goes unnoticed. This is unsettling when we consider the important role assessing the evidence plays in establishing the facts in criminal proceedings.

The question of whether the assessment of testimonies really is as superficial as the analysis of court decisions suggests remains unanswered. One can argue that in essence it is: the evaluations are generalised and this is reflected in the justifications, in which judges call upon only a few well-rehearsed assessment criteria. One might also argue that assessment of testimony is thorough at the stage of examining witnesses or before deciding the ruling, and only the written reason for judgment, which is usually compiled a few hours after the ruling has been announced, is standardised and general. That the written reasons for judgment do not reflect the true process of assessing testimony (though they should) cannot be ruled out. There are a number of possible reasons for this disparity. First of all, routine. Judges prepare their reasons for judgment in a schematic and generic manner that is reproduced in case after case. Secondly, it may be the result of confirmation bias (Nisbett, Ross, 1980; Rosenthal, 1994; Synder, Swan, 1978). Since the court decides the guilt or innocence of the accused at the stage of delivering a ruling, there may be a tendency when writing the reason for judgment to interpret the evidence to be assessed in accordance with that decision as to responsibility, or absence of responsibility, for the crime. In other words, if the court has delivered a guilty verdict it may value the reliability of evidence (or of aspects of the evidence, such as the testimony of a witness) that supports the guilt of the accused and overlook evidence (or aspects of the evidence) that contradicts the thesis that the accused is responsible for the crime. Thirdly and finally, the superficiality of evaluations could result from the fact that judges are guided by their intuition in assessing witness testimony and only *post factum*, when preparing the reasons for judgment, do they support their assessment of reliability by referring to recognisable evaluation criteria. No matter the reasons for the aforementioned generality of assessments of the reliability of witness testimony in the written reasons for judgment, it has real consequences. If one of the parties to the proceedings decides to lodge an appeal, it is obliged to set out the charges against the disputed ruling. If, for example, a party argues on appeal that the judge assessed the evidence incorrectly (by giving credence to the testimony of a witness who, in the opinion of the appellant, was unreliable) it should be in a position to demonstrate the court's errors of reasoning and assessment. A precise evaluation reproduced in writing ena-

bles the appeal court to follow the assessment made by the lower court. It is rare for courts to hold evidential proceedings in appeals cases (especially full evidential proceedings), and they are particularly unlikely to do so where this has already been done by the lower court.

It would seem that when judges finally acknowledge testimony as unreliable (or partially unreliable), they take a slightly more thorough approach to the assessment of testimony. Given that almost 80% of the testimonies of all witnesses assessed were acknowledged as reliable, the impression can be gained that to acknowledge testimony as reliable is something of a norm in Polish criminal proceedings. Yet when courts deny the reliability of testimony, they advance more legal criteria to explain it than when they acknowledge testimony as reliable. In these cases the courts produce assessments of testimony without referring to evaluation criteria significantly more rarely (and in overall terms exceptionally rarely) than they do when assessing testimony deemed reliable. Not only do courts give more objective criteria when assessing unreliable testimony, they also give more subjective criteria. It should be emphasised, though, that there was no statistically significant difference between reliable and unreliable testimony in the number of witness-statement assessments, in which subjective criteria outnumbered objective criteria. It is interesting that such an important and significant difference affected testimony assessed according to the same number of criteria of both types – a situation that occurred more often in the case of testimony deemed unreliable. This would appear to confirm the view expressed by Płachta (1985) that courts tend to justify the unreliability of testimony by subjective criteria and its reliability by objective criteria. In Płachta's view (1985), this results from judges studying the contents of the evidence – and finding sufficient grounds there to support their assessments – in cases where they deem testimony to be reliable. Yet when they doubt the reliability of testimony they search for additional arguments to undermine the evidential resource under analysis, which in this case is, in a sense, “external” to the content of the evidence and concerns its source.

The results obtained in this study allow interesting conclusions to be drawn about the influence of certain characteristics both of criminal proceedings, and of the witnesses themselves, on the evaluation of testimony. Judges employ significantly fewer criteria to evaluate testimony lodged in cases involving high-severity crimes when compared to cases involving medium-severity crimes and low-severity crimes. Moreover, this affects the number of both objective

and subjective criteria employed to assess testimony given in high-severity proceedings. Given that courts trying crimes such as sexual abuse of a minor, fraud or robbery are expected to be especially scrupulous in evaluating evidence, this may be of particular concern. It can only be hoped (though the present research provides no evidence to sustain it) that the more severe the crime is, the more the courts will rely on means of proof other than personal evidence. Where high-severity crimes are concerned, the actions of the law-enforcement authorities are perhaps directed with greater intensity towards searching for tangible evidence that therefore then dominates the evidential hearings in court.

The frequency and type of assessment criteria invoked by the courts differed according to their size. Objective criteria were mentioned significantly more often, and subjective criteria more rarely, in evaluations of testimony at the District Court for Kraków-Nowa Huta (known as a large court because it is in a place that is also the seat of a provincial court) than at Brzesko District Court (known as a small court because it is not in a place where a provincial court also sits). Assessments of testimony by judges in which subjective criteria outnumbered objective criteria, and assessments of testimony in which the number of criteria of each type was equal, were found significantly more often at the “small” district court than at the “large” court. It is difficult to state categorically why the larger court exercised more reserve with regard to subjective assessment criteria when compared to the smaller court.

It turns out that the role of the witness has a significant influence on the number and type of criteria invoked in evaluations of testimony. Firstly, judges draw on a greater number of criteria when assessing the testimony of victims than when assessing the testimony of disinterested witnesses. Furthermore, they invoke objective criteria more often than subjective criteria in assessments of victims compared to assessments of testimony given by disinterested witnesses. Finally, the results suggest that the testimony of victims is more often regarded as reliable than the testimony of disinterested witnesses. In other words, the courts assess the testimony of victims more thoroughly and they are more likely to justify assessments of the testimony of victims by reference to objective criteria (and therefore by reference to the content of the testimony). What is more, they regard the testimony of victims as reliable more often than the testimony of disinterested witnesses. It would therefore appear that judges trust the testimony of victims more than that of disinterested witnesses. At the same time, they are more

exact in determining whether victims of crime really do deserve this trust. Furthermore, the judges justify these assessments by deploying more arguments that are concrete and tangible than arguments of a personal nature. Could the judges believe that victims’ motivation to testify, including giving testimony with specific content, is, in the majority of cases, dictated by an interest (considered suspicious) in securing a particular outcome (usually the conviction of the accused)? Judges may perceive this as automatically casting a shadow of doubt over the reliability of victims. It is perhaps exactly for this reason that judges are more inclined to justify the reliability of the testimony of victims by invoking criteria which in their opinion are more convincing.

A number of cautious conclusions can be drawn in the light of these results with regard to the influence of age on assessments of witness testimony. Firstly, it was noted that the older the witness, the fewer criteria judges invoke in their evaluations of testimony. This relationship was found to be weak, but it suggests that the courts make more thorough evaluations of the testimony of younger witnesses, including children. Numerous studies of children’s testimony (there is a literature review in Principe, Greenhoot, Ceci, 2014) have indicated that children – especially younger ones – supply less information than adults and older children (Pipe, Lamb, Orbach, Esplin, 2004), give inaccurate details of the age, weight and height of criminals more often (Pozzulo, 2007) and are more suggestible than adults (Melnyk, Crossman, Scullin, 2007). It is possible that judges are aware of the limitations of children’s memories, and of their testimony, and therefore take a more searching approach to their witness statements. It is a matter of concern that a similar effort is not made with the testimony of elderly witnesses, whose cognitive functioning differs from that of adults (Mueller-Johnson, Ceci, 2007). However the findings of researchers who study the memory and testimony of elderly people are considerably fresher, sharper and more acute than the testimony of the children we met above – but this knowledge has so far not percolated through the judiciary. The observations regarding the influence of age on whether testimony is seen as reliable or unreliable are slightly more disturbing. It turns out that the testimony of witnesses aged fifteen or under was regarded as unreliable more often than in the remaining categories of witnesses deemed unreliable. Yet because not all of the requirements of the statistical analysis were met, only very cautious conclusions can be drawn. Professionals in the justice system have often found that the testimony of children is less trustworthy than that of adults (Benton, Ross,

Bradshaw, Thomas, Bradshaw, 2006; Granhag et al., 2005). Suspicion of the testimony of this category of witnesses prevailed long in a number of jurisdictions and perhaps continues to do so (Bala, Ramakrishnan, Lindsay, Lee, 2005). It is not impossible that this suspicion will also filter through into the Polish criminal justice system.

It was also noted that courts invoke a significantly greater number of testimony-assessment criteria when they are evaluating the testimony of witnesses examined in the presence of a psychologist (as an expert witness) than when there is no psychologist present. In accordance with article 192 § 2 of the Code of Criminal Procedure, if there exists doubt as to the mental state, degree of intellectual development, or capacity to receive or reproduce perceptions of a witness, the court or the public prosecutor can order their evidence to be heard in the presence of a doctor or psychologist, and the witness cannot contest the order. Pursuant to article 192 § 2 of the Code of Criminal Procedure, moreover, witnesses can be subjected to psychological examination. In practice, therefore, hearing evidence in the presence of an expert witness also involves writing a forensic-psychological opinion on the psychological validity of the testimony given. The greater number of testimony-assessment criteria invoked in the case of witnesses whose evidence was also evaluated by a psychologist is probably due to the fact that judges draw on arguments in psychological reports to support their own, legal, evaluation of the evidence from the witness testimony. Many criteria of legal evaluation are, after all, consistent with the criteria for assessing the psychological validity of testimony in the most widely-used tool for this purpose: Statement Validity Analysis (SVA), which we have already encountered above. SVA is made up of a number of components, of which the most important is based on Criteria-Based Content Analysis (CBCA), which contains nineteen content criteria for the validity of testimony. One example of a content criterion is logical structure, which is satisfied when the testimony makes fundamental sense, is coherent and logical, is not self contradictory and meets the legal criteria of logicity and internal consistency (Memon et al., 2003). Further CBCA criteria – quantity of details and unstructured production – correspond with the legal criteria of the degree of detail in testimony and its spontaneity.

Interesting results were noted with respect to the testimony of witnesses related to the accused. Judges give a greater number of criteria in their justifications when assessing the testimony of witnesses who are closely related to the accused (in the sense of people entitled to refuse to testify according to article 182 of

the Code of Criminal Procedure), when compared to witnesses who are not. More than this, judges drew on no assessment criteria whatever in their reasons for judgment significantly more rarely in the case of the testimony of people who are closely related to the accused when compared to witnesses who are not. With respect to this category of witnesses the courts invoked a greater number of both objective and subjective criteria. It can be assumed that judges are more scrupulous when assessing the testimony of witnesses who are closely related to the accused. It is interesting that – irrespective of the thoroughness of the assessment – the testimony of those closely related to the accused was regarded as unreliable significantly more often than the testimony of those not closely related to the accused. This might suggest a certain prejudice on the part of the courts against witnesses who are closely related to the accused, though there is no doubt that the courts should approach all forms of evidence objectively and without bias (Płachta, 1985). It would be naive, however, to believe that judges are not prone to various forms of prejudice and that they do not influence decisions taken during criminal proceedings – especially where these decisions concern the person of the accused (Rachlinski, Johnson, Wistrich, Guthrie, 2009; Wistrich, Rachlinski, 2017). Nevertheless, given that it is possible for judges to be partial because of certain characteristics or attributes of the accused, it is surely also possible that they are prone to prejudice with regard to witnesses (and therefore this prejudice can indirectly influence assessments of the guilt or innocence of the accused). It must of course be borne in mind that judges already apply many of the measures that the literature regards as minimising their influence on their own decisions: they are well educated and well trained, they possess considerable experience of taking difficult decisions and they are motivated to ensure that they reach just verdicts. They take responsibility for their judgments, which they take in public, their decisions are subject to appeal, and career advancement is dependent on an examination of the body of rulings they have made in the past (Wistrich, Rachlinski, 2017).

The legal and psychological literature informs us that evidence from witness testimony (and first of all errors in testimony and false testimony) is one of the principle causes of the problem of judicial mistakes, which is faced by legal systems the world over (Huff, 1987; Rattner, 1988, for example). The superficiality of courts' assessments of witness testimony, which has been identified in the empirical research set out in this study, can, to a certain extent, explain why unjust convictions are so often the result of false or untrue testi-

mony (Scheck, Neufeld, Dwyer, 2000, for example). This should, however, make judges sensitive to the importance of thorough and detailed evaluation of testimony. The legal nature of assessments of testimony (and other evidence brought forward in proceedings) does not mean that they can be produced as if in isolation from contemporary psychological knowledge. There is no way to eliminate the mistakes made by eyewitnesses to crimes, which will always be with us. It should nevertheless be possible to restrict their influence on criminal proceedings. The skilful application in practice of the knowledge of judges can go a long way to achieving this.

Acknowledgements

The research was funded by the National Science Centre in Poland (grant number: 2015/19/D/HS5/01241).

References

- Bala, N., Ramakrishnan, K., Lindsay, R., Lee, K. (2005). Judicial assessment of the credibility of child witnesses. *Alberta Law Review*, 42(4), 995–1017.
- Benton, T. R., Ross, D. F., Bradshaw, E., Thomas, N. W., Bradshaw, G. S. (2006). Eyewitness memory is still not a common sense: comparing jurors, judges and law enforcement to eyewitness experts. *Applied Cognitive Psychology*, 20(10), 115–129. doi: 10.1002/acp.1171.
- Granhag, P. A., Strömwall, L. A., Hartwig, M. (2005). Eyewitness testimony: tracing the beliefs of Swedish professionals. *Behavioral Sciences & the Law*, 23(5), 709–727. doi: 10.1002/bsl.670.
- Gruza, E. (2003). *Ocena wiarygodności zeznań świadków w procesie karnym. Problematyka kryminalistyczna*. Kraków: Zakamycze.
- Gruza, E. (2014). Stan dyscypliny psychologia i prawo okiem kryminalistyka. (In) E. Habzda-Siwiek, J. Kabzińska (Eds.), *Psychologia i prawo. Między teorią a praktyką* (pp. 31–45). Sopot: GWP.
- Huff, C. R. (1987). Wrongful conviction: societal tolerance of injustice. (In) J. L. Miller, M. Lewis (Eds.), *Research in social problems and public policy* (pp. 99–115). Greenwich, CT: JAI.
- Maciejski, M. (2009). *Psychologiczna analiza sposobów przesłuchania świadków i reguł oceny ich zeznań w praktyce sędziowskiej a stopień przypisywanej im wiarygodności* (unpublished doctoral thesis). Katowice: Uniwersytet Śląski.
- Magnussen, S., Wise, R. A., Raja, A. Q., Safer, M. A., Pawlenko, N., Stridbeck, U. (2008). What judges know about eyewitness testimony: A comparison of Norwegian and US judges. *Psychology, Crime & Law*, 14(3), 177–188. doi: 10.1080/10683160701580099.
- Melnyk, L., Crossman, A. M., Scullin, M. H. (2007). The suggestibility of children's testimony. (In) M. P. Toglia, J. D. Read, R. C. L. Lindsay (Eds.), *Handbook of eyewitness psychology: Volume I: Memory for events* (pp. 401–427). Mahwah, NJ: Earlbaum Associates.
- Memon, A., Vrij, A., Bull, R. (2003). *Prawo i psychologia. Wiarygodność zeznań i materiału dowodowego*. Gdańsk: GWP.
- Mueller-Johnson, K., Ceci, S. J. (2007). The elderly eyewitness: A review and prospectus. (In) M. P. Toglia, J. D. Read, R. C. L. Lindsay (Eds.), *Handbook of eyewitness psychology: Volume I: Memory for events* (pp. 577–603). Mahwah, NJ: Earlbaum Associates.
- Nisbett, R. E. Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.
- Piekarska-Drażek, M. (2004). *Zmiana zeznań świadków w procesie karnym* (unpublished doctoral thesis). Warszawa: Uniwersytet Warszawski.
- Pipe, M.-E., Lamb, M. E., Orbach, Y., Esplin, P. W. (2007). Recent research on children's testimony about experienced and witnessed events. *Developmental Review*, 24(4), 440–468. doi: 10.1016/j.dr.2004.08.006.
- Płachta, M. (1985). Wykorzystanie i ocena zeznań świadka w procesie karnym. *Zeszyty Naukowe UJ*, 116, 130–132.
- de Poot, C. J., Bokhorst, R. J., van Koppen, P. J., Muller, E. R. (2004). *Rechercheportret: over dilemma's in de opsporing*. Alphen aan den Rijn: Kluwer.
- Pozullo, J. (2007). Person description and identification by child witnesses. (In) R. C. L. Lindsay, D. F. Ross, J. D. Read, M. P. Toglia (Eds.), *The handbook of eyewitness psychology. Volume II: Memory for people* (pp. 283–307). Mahwah, NJ: LEA Publishers.
- Principe, G. F., Greenhoot, A. F., Ceci, S. J. (2014). Young children's eyewitness memory. (In) T. J. Perfect, D. S. Lindsay (Eds.), *The Sage handbook of applied memory* (pp. 633–653). London: Sage Publications.
- Rachlinski, J. J., Johnson, S., Wistrich, A. J., Guthrie, C. (2009). Does unconscious racial bias affect trial judges? *Cornell Law Faculty Publications*. Paper 786.
- Rattner, A. (1988). Convicted but innocent: wrongful convictions and the criminal justice system. *Law and Human Behavior*, 12(3), 283–293. doi: 10/1007/BF01044385.
- Rosenthal, R. (1994). Interpersonal expectancy effects: a 30 year perspective. *Current Directions in Psychological Science*, 3(6), 176–179. doi: 10.1111/1467-8721.ep10770698.
- Scheck, B., Neufeld, P., Dwyer, J. (2000). *Actual innocence: five days to execution and other dispatches from the wrongly convicted*. Garden City, NY: Doubleday.

23. Semmler, C., Brewer, N., Douglass, A. B. (2011). Jurors believe eyewitnesses. (In) B. L. Cutler (Ed.), *Conviction of the innocent: Lessons from psychological research* (pp. 185–209). Washington, DC: American Psychological Association.
24. Sporer, S. L. (2008). Lessons from the origins of eyewitness testimony research in Europe. *Journal of Applied Cognitive Psychology*, 22(6), 737–757. doi: 10.1002/acp.1479.
25. Synder, M., Swann, W. B. (1978). Behavioral confirmation in social interaction: From social perception to social reality. *Journal of Experimental and Social Psychology*, 14, 148–162. doi: 10.1016/0022-1031(78)90021-5.
26. Vallas, J. (2011). A survey of federal and state standards for the admission of expert testimony on the reliability of eyewitnesses. *American Journal of Criminal Law*, 39(1), 97–146.
27. Waltoś, S. (1985). *Świadek w procesie sądowym*. Warszawa: Wydawnictwo Prawnicze.
28. Waltoś, S., Hofmański, P. (2013). *Proces karny. Zarys systemu*. Warszawa: LexisNexis.
29. Wise, R. S., Dauphinais, K. A., Safer, M. A. (2007). A tripartite solution to eyewitness error. *Journal of Criminal Law & Criminology*, 97(3), 807–872.
30. Wise, R. A., Safer, M. A. (2004). What US judges know and believe about eyewitness testimony. *Applied Cognitive Psychology*, 18(4), 427–443. doi: 10.1002/acp.993.
31. Wistrich, A. J., Rachlinski, J. J. (2017). Implicit bias in judicial decision making how it affects judgment and what judges can do about it. *Cornell Legal Studies Research Paper* No. 17–16. doi: 10.2139/ssrn.2934295.

Corresponding author

Dr Joanna Kabzińska
SWPS University of Social Sciences and Humanities,
Warsaw
Katowice Faculty of Psychology
ul. Techników 9
PL 40-326 Katowice
e-mail: jkabzinska@swps.edu.pl

OCENA WIARYGODNOŚCI ZEZNAŃ ŚWIADKÓW W POSTĘPOWANIU KARNYM – WSTĘPNE WYNIKI BADAŃ EMPIRYCZNYCH

Wprowadzenie

Dokonywanie ustaleń faktycznych w postępowaniu sądowym w każdym przypadku wymaga oceny całości dowodów zgromadzonych w sprawie i uwzględnienia ich wzajemnego powiązania. Na straży takiego modelu rozstrzygnięcia stoi wyrażona w art. 7 kodeksu postępowania karnego zasada swobodnej oceny dowodów nakazująca organom procesowym kształtowanie swojego przekonania na podstawie wszystkich przeprowadzonych dowodów z uwzględnieniem zasad prawidłowego rozumowania, wskazań wiedzy i doświadczenia życiowego. Ocenę dowodu zgodną ze wskazaniem nauki należy rozumieć jako ocenę pozostającą w zgodzie z najnowszymi i wiarygodnymi wynikami badań naukowych. Chodzi tu w szczególności o posługiwanie się metodą naukową oraz wykorzystywanie osiągnięć nauki (Waltoś, Hofmański, 2013). W procesie swobodnej oceny dowodów z zeznań świadków szczególnie przydatna jest wiedza z zakresu psychologii zeznań świadków. Choć podstawowym narzędziem oceny dowodów przez organy procesowe są niewątpliwie prawa logiki, to znajomość psychologii jest dla prawnika niezbędnym instrumentem pracy, gwarancją niezawisłości i trafności wydawanych orzeczeń (Gruza, 2014). Okazją do zastosowania wiedzy psychologicznej w praktyce przez organy procesowe jest dokonywana przez nie ocena wiarygodności dowodu z zeznań świadka. W literaturze psychologicznej i kryminalistycznej odnaleźć można doniesienia dotyczące kryteriów takiej oceny (np. Gruza, 2003; Maciejewski, 2009; Płachta, 1985). Mimo to kryteria tej oceny nie są wśród praktyków rozpowszechnione, a nierzadko nie są im w ogóle znane (Gruza, 2014). Wskazane w orzecznictwie przykłady, jakimi kryteriami kierować się (a części – jakimi kierować się nie należy), trudno uznać za wyczerpujące (np. wyrok SN z 11 stycznia 1996, II KRN 178/95).

Stosunkowo niewiele wiadomo o wpływie dorobku psychologii zeznań świadków na praktykę funkcjonowania wymiaru sprawiedliwości. Wiedza uczestników postępowania karnego z psychologii zeznań świadków była dotychczas badana w drodze badań ankietowych. Badania z udziałem sędziów amerykańskich (Wise, Safer, 2003), ale także norweskich (Magnussen i in. 2008) i szwedzkich (Granhag, Strömwall, Hartwig, 2005) wskazują, że wiedzą oni niewiele o czynnikach i procesach wpływających na percepcję i pamięć naocznych świadków. Badania pokazują tymczasem, że dysponowanie przez sędziów wiedzą psychologiczną wpływa pozytywnie na ich praktykę orzeczniczą (Wise,

Safer, 2004). Niejednokrotnie zresztą edukacja sędziów i innych uczestników postępowania karnego wskazywana jest jako istotny element służący minimalizacji problemu nieprawdziwych czy fałszywych zeznań w procesie karnym (np. Wise, Dauphinais, Safer, 2007). Podobne próby badawcze zmierzające do ustalenia wiedzy psychologicznej polskich zawodowych uczestników postępowania karnego i jej wpływu na postępowanie podejmowane były w znikomym stopniu (por. Piekarska-Drażek, 2004). Jednak sam fakt stwierdzenia posiadanej przez prawników wiedzy z określonej dziedziny nie jest równoznaczny z umiejętnością jej wykorzystania.

Sporer (2008) wskazywał, że dotychczasowe badania w obszarze psychologii zeznań świadków niemal zupełnie pominęły kwestię tego, jak relacje świadków są oceniane. Zaleca jednocześnie skoncentrowanie uwagi badaczy na tym, w jaki sposób organy procesowe oceniają zeznania świadków (Sporer, 2008). W tym celu podjęto badania dotyczące oceny wiarygodności ich zeznań skoncentrowane na badaniu praktyki dokonywania tej oceny. Dokonano tego na podstawie analizy akt sądowych w postępowaniach karnych. Postawiono następujące pytania badawcze:

1. Jakimi kryteriami oceny wiarygodności posługują się w praktyce sędziowie podczas oceny zeznań świadków?
2. Jakie czynniki wpływają na liczbę i rodzaj stosowanych kryteriów wiarygodności zeznań?
3. Jakie czynniki wpływają na ostateczny rezultat oceny zeznań świadków?

Materiały i metoda

Analizą objęto sprawy karne z dwóch sądów w obszarze właściwości miejscowej Sądu Apelacyjnego w Krakowie – z Sądu Rejonowego w Brzesku i z Sądu Rejonowego dla Krakowa-Nowej Huty¹. Do władz obu sądów wystosowano prośbę o dostęp do akt 54 spraw karnych spełniających poniższe kryteria:

- a) sprawy prawomocnie zakończone w latach 2013–2015;
- b) sprawy, w których przesłuchano przynajmniej jednego świadka;

¹ Docelowo badania zostaną przeprowadzone w 8 sądach rejonowych w Polsce znajdujących się w obszarze właściwości miejscowej 4 sądów apelacyjnych w Krakowie, Poznaniu, Lublinie i Gdańsku.

- c) sprawy, w których sporządzono pisemne uzasadnienie orzeczenia przez sąd I instancji;
- d) sprawy należące do kategorii lekkich, średnich i ciężkich przestępstw².

Występki spełniające kryterium powagi przestępstwa wybrano w ten sposób, że z 22 rozdziałów części szczególnej kodeksu karnego wyselekcjonowano te rozdziały, w związku z którymi:

- a) w 2015 roku doszło do największej liczby skazań, nie mniej jednak niż 1500 osób;
- b) wyeliminowano te, w związku z którymi postępowanie dowodowe opiera się w większym stopniu na innych niż zeznania świadków środkach dowodowych.

W wyniku zastosowania tych kryteriów wyselekcjonowano 7 rozdziałów kodeksu karnego. Z każdego wybrano maksymalnie po 3 występkę, w efekcie których w 2015 roku doszło do największej liczby skazań (przy czym nie mniej niż 500 osób). W rezultacie tej analizy wybrano 15 występków. Badania pilotażowe dodatkowo doprowadziły do wyeliminowania z tej puli 3 z nich. W efekcie wybrano po 4 kwalifikacje prawne czynów, które spełniały wspomniane kryterium powagi przestępstwa, a były to – pośród przestępstw ciężkich: a) obcowanie płciowe z małoletnim poniżej 15 roku życia, b) łapownictwo czynne, c) oszustwo, d) rozbój; pośród średnich: a) średni uszczerbek na zdrowiu, b) przestępstwo zmuszania w celu wymuszenia zwrotu wiarygodności, c) przestępstwo znęcania, d) przestępstwo kradzieży; wreszcie pośród lekkich: a) lekki uszczerbek na zdrowiu, b) przestępstwo gróźb karalnych, c) naruszenie nietykalności cielesnej funkcjonariusza publicznego, d) znieważenie funkcjonariusza publicznego.

Analiza akt spraw sądowych ze szczególnym uwzględnieniem pisemnego uzasadnienia orzeczenia sądu została przeprowadzona na podstawie kwestionariusza składającego się z 3 części – dotyczących sprawy, świadka i oceny zeznania. Kryteria oceny wiarygodności zostały wybrane na podstawie dostępnej literatury (Gruza, 2003; Maciejski, 2009; Płachta, 1985), orzecznictwa polskich sądów apelacyjnych i Sądu Najwyższego oraz badań pilotażowych kwestionariusza ($N = 28$). Podzielono je na kryteria przedmiotowe – odnoszące się do treści zeznań i podmiotowe – odnoszące się do osoby świadka. Informacje na temat zamieszczonych w kwestionariuszu prawnych kryteriów oceny zeznań wraz z wyjaśnieniem zawiera tabela 1.

² Kryterium powagi przestępstwa stanowiło zagrożenie karą pozbawienia wolności za dane przestępstwo przypisane w kodeksie karnym. Przestępstwa lekkie to takie, których górna granica ustawowego zagrożenia wynosiła 3 lata pozbawienia wolności, średnie – dolna granica ustawowego zagrożenia karą pozbawienia wolności wynosiła ponad 3, ale mniej niż 5 lat oraz ciężkie – w których górna granica ustawowego zagrożenia karą pozbawienia wolności wynosiła ponad 5 lat.

Próba

Analizą objęto 90 spraw karnych z dwóch sądów w obszarze właściwości miejscowej Sądu Apelacyjnego w Krakowie – Sądu Rejonowego w Brzesku ($N = 44$) i Sądu Rejonowego dla Krakowa-Nowej Huty ($N = 46$). Szczegółowe informacje na temat struktury próby zawiera tabela 2.

W objętych analizą sprawach przesłuchano 751 świadków (w SR w Brzesku, $N = 361$, w SR dla Krakowa-Nowej Huty, $N = 390$). Próba składała się z 292 kobiet (38,9%) i 459 mężczyzn (61,1%). Przesłuchano zaledwie 2 świadków poniżej 15 roku życia i 70 świadków powyżej 65 roku życia. Szczegółowe informacje na temat struktury próby w odniesieniu do świadków zawiera tabela 3.

Wyniki

Ogólne wyniki dotyczące oceny zeznań przez sąd

Średnia liczba świadków przesłuchanych w analizowanych sprawach wyniosła 8,34 ($SD = 6,15$). W obu sądach nie dokonano oceny zeznań 104 świadków spośród wszystkich przesłuchanych w postępowaniu. Powody odmowy oceny zeznań były różne: dominowało uzasadnienie, że świadek nie posiadał informacji istotnych dla sprawy, choć były przypadki, w których sąd nie podjął w ogóle kwestii przyczyn niedokonania oceny³. W większości przypadków zeznania świadków zostały uznane za wiarygodne (w SR w Brzesku $N = 229$; w SR dla Krakowa-Nowej Huty $N = 284$), natomiast w 134 przypadkach – za niewiarygodne lub częściowo niewiarygodne (w SR w Brzesku $N = 70$; w SR dla Krakowa-Nowej Huty $N = 64$). Sądy w objętych analizą sprawach powołały w uzasadnieniach wyroków 1723 kryteria oceny zeznań. Średnio przywołano 2,65 ($SD = 1,7$) kryterium w ocenie jednego zeznania. W ocenie dominowały kryteria przedmiotowe, a zatem te odnoszące się do treści zeznania. Zastosowano ogółem 1472 kryteria o charakterze przedmiotowym ($M = 2,26$; $SD = 1,49$) i 251 kryteriów podmiotowych, czyli dotyczących osoby świadka ($M = 0,39$; $SD = 2,26$). Nie w każdym z analizowanych zeznań przywołano określony typ kryteriów. Kryteria przedmiotowe zostały przywołane w 599 ocenianych zeznaniach ($M = 2,46$; $SD = 1,39$), natomiast podmiotowe – w zaledwie 196 ($M = 1,28$; $SD = 0,55$). Częstotliwość przywoływania poszczególnych kryteriów o charakterze przedmiotowym i podmiotowym w obu objętych analizą sądach zawierają tabele 4 i 5.

Ogólnie w obu analizowanych sądach (analizowane łącznie) spośród przywołanych w uzasadnieniu wyro-

³ Dalsze analizy zostały przeprowadzone z wyłączeniem tychże 104 przypadków.

ku kryteriów przedmiotowych najczęściej posługiwano się (w przypadku uznania zeznań za wiarygodne) kryterium zgodności zeznania z innymi dowodami w sprawie ($N = 435$), wewnętrznej spójności złożonych zeznań ($N = 192$) oraz logiczności zeznań ($N = 176$). W Sądzie Rejonowym dla Krakowa-Nowej Huty tak właśnie prezentowała się trójka najczęściej stosowanych kryteriów przedmiotowych (przy czym drugim najczęściej stosowanym kryterium była logiczność zeznań, a wewnętrzna spójność – trzecim). W Sądzie Rejonowym w Brzesku sędziowie najczęściej przywoływali kryterium zgodności zeznań z innymi dowodami, wewnętrznej spójności i – na trzecim miejscu – obiektywizmu zeznań. Gdy zeznania uznawane zostały za niewiarygodne (lub częściowo niewiarygodne), najczęściej przywoływanymi kryteriami braku wiarygodności były: niezgodność ze zgromadzonym w sprawie materiałem dowodowym, brak spójności zeznań i ogólnikowość zeznań.

W obu (analizowanych łącznie) sądach spośród przywołanych w uzasadnieniu wyroku kryteriów podmiotowych najczęściej posługiwano się (w przypadku uznania zeznań za wiarygodne) kryterium innej wiarygodnej motywacji do złożenia zeznań ($N = 66$), wiarygodnej motywacji wynikającej z braku powiązania z oskarżonym ($N = 43$) i spontaniczności w składaniu zeznań ($N = 11$). Wystąpiły jednak różnice między poszczególnymi sądami. O ile wspomniane dwa kryteria motywacyjne były najczęściej przywoływane w uzasadnieniach w obu sądach, o tyle na trzecim miejscu w Sądzie Rejonowym dla Krakowa-Nowej Huty jako najczęściej używane kryterium znalazło się „inne kryterium wiarygodności”, a w Brzesku – wspomniana spontaniczność w sposobie złożenia zeznań.

Wybrane zmienne niezależne a liczba kryteriów oceny zeznań przywołanych w uzasadnieniu wyroku

Analiza wyników wskazuje na istnienie statystycznie istotnej zależności między pewnymi właściwościami postępowania oraz charakterystykami świadka a liczbą kryteriów oceny wiarygodności zeznań przywołanych przez sędziów w uzasadnieniu (Tabela 6).

Analiza statystyczna przy użyciu nieparametrycznego testu Kruskala-Wallisa wskazuje na istnienie statystycznie istotnej różnicy między liczbą przywołanych kryteriów a powagą przestępstwa będącego przedmiotem sprawy ($H = 15,32, p < 0,05$). W celu sprawdzenia które z różnic są statystycznie istotne przeprowadzono serię porównań z wykorzystaniem testu *U* Manna-Whitneya. Sędziowie używają mniej kryteriów oceny wiarygodności do analizy zeznań w sprawach z kategorii ciężkich w porównaniu ze sprawami dotyczącymi przestępstw średnich ($Z = -2,57, p < 0,05$) oraz lekkich ($Z = -3,92, p < 0,05$). Stwierdzono również statystycznie istotną zależność między liczbą przywołanych kryteriów

oceny zeznań a surowością orzeczonej przez sąd kary ($H = 19,26, p < 0,05$). Dalsza analiza z wykorzystaniem testu *U* Manna-Whitneya wskazuje, że sędziowie powołują statystycznie istotnie mniej kryteriów, gdy wymierzona kara jest surowsza od tej wnioskowanej przez prokuratora niż wówczas, gdy wymierzona kara jest tożsama z wnioskowaną przez prokuratora ($Z = -4,315, p < 0,05$) oraz łagodniejsza od wnioskowanej przez prokuratora ($Z = -3,78, p < 0,05$).

Wyniki wskazują na statystycznie istotną zależność między rolą świadka w postępowaniu, wiekiem świadka i związkiem świadka z oskarżonym a liczbą przywołanych w uzasadnieniu kryteriów oceny. Sędziowie przywołują więcej kryteriów w ocenie zeznań pokrzywdzonych niż świadków ($Z = -7,11, p < 0,05$). Ponadto analiza korelacji *r*-Pearsona wskazuje na słaby, lecz statystycznie istotny ujemny związek pomiędzy wiekiem świadka a liczbą przywoływanych kryteriów ($r = -0,1, p < 0,05$). Co więcej, wyniki przeprowadzonych analiz sugerują statystycznie istotną różnicę w liczbie przywołanych przez sąd kryteriów oceny wiarygodności zeznań w zależności od związku między świadkiem a oskarżonym. Sędziowie używają więcej kryteriów wiarygodności do oceny zeznań osób najbliższych oskarżonemu niż osób mu obcych ($Z = -4,32, p < 0,05$). Udział biegłego psychologa podczas przesłuchania świadka również ma wpływ na liczbę przywoływanych w uzasadnieniach kryteriów oceny zeznań. Sędziowie używają ich istotnie więcej do oceny zeznań świadków, w przesłuchaniu których brał udział biegły psycholog (i sporządził opinię dotyczącą psychologicznej prawdziwości ich zeznań) niż gdy nie był on obecny przy przesłuchaniu ($Z = -3,24, p < 0,05$). Wreszcie, stwierdzono istotną statystycznie różnicę w liczbie stosowanych kryteriów oceny wiarygodności w zależności od ostatecznej oceny zeznania dokonanej przez sąd. Sędziowie przywołują w uzasadnieniach więcej kryteriów, gdy uznają zeznania za niewiarygodne (lub częściowo niewiarygodne) niż wiarygodne ($Z = -4,96, p < 0,05$).

Wybrane zmienne niezależne a rodzaj kryteriów oceny zeznań przywołanych w uzasadnieniu wyroku

Przedmiotem analizy uczyniono również rodzaj kryteriów przywoływanych przez sędziów w uzasadnieniu wyroku (przedmiotowe vs. podmiotowe). W pierwszej kolejności porównano, czy wybrane zmienne zależne wpływają na stosunek kryteriów podmiotowych do przedmiotowych przywołanych w ocenie zeznania (Tabela 7).

Analiza wykazała, że wielkość sądu (duży sąd vs. mały sąd), w którym orzekano, wpływa na wybór kryteriów przez sędziów uzasadniających ocenę zeznań⁴.

⁴ Analizę przeprowadzono w dwóch sądach – dużym sądzie rejonowym (SR dla Krakowa-Nowej Huty), rozumianym

W małych sądach istotnie częściej pojawiały się oceny zeznań świadków uzasadnione z przewagą kryteriów o charakterze podmiotowym nad przedmiotowymi (6,6%) niż w sądzie dużym (1,1%). Jednocześnie w małych sądach istotnie częściej zdarzały się oceny zeznań, w których liczba przywołanych kryteriów podmiotowych i przedmiotowych była równa (11,9%) niż w dużych sądach (3,4%, $\chi^2(3) = 32,29$, $p < 0,05$).

W przypadku lekkich występzków istotnie częściej oceniano złożone zeznania z przewagą kryteriów przedmiotowych nad podmiotowymi (90,1%) w porównaniu z zeznaniami złożonymi w przypadku średnich (78,7%) i lekkich (84,8%) występzków. Natomiast w przypadku występzków średniej wagi sędziowie dwukrotnie częściej uzasadniali swą ocenę, przywołując tę samą liczbę kryteriów przedmiotowych co podmiotowych niż w przypadku pozostałych kategorii występzków ($\chi^2(6) = 16,24$, $p < 0,05$).

Stwierdzono również, że w przypadku zeznań pokrzywdzonych istotnie częściej uzasadniano ocenę ich zeznań z przewagą kryteriów przedmiotowych nad podmiotowymi (95,9%) w porównaniu z zeznaniami świadków postronnych (79,9%, $\chi^2(3) = 21,21$, $p < 0,05$). W przypadku świadków powiązanych rodzinnie z oskarżonym istotnie rzadziej zdarzały się takie oceny zeznań, w których sąd nie przywołał żadnych kryteriów (0,9%) w porównaniu z oceną świadków niepowiązanych z oskarżonym (6,3%). Ponadto w przypadku świadków powiązanych istotnie częściej oceniano ich zeznania, używając równej liczby kryteriów przedmiotowych i podmiotowych (13,2%) niż niepowiązanych (6,3%, $\chi^2(3) = 10,69$, $p < 0,05$).

Wreszcie, na stosunek kryteriów podmiotowych do przedmiotowych w ocenie zeznań ma wpływ także finalna ocena złożonych zeznań jako wiarygodnych lub niewiarygodnych. Istotnie częściej dochodziło do oceny zeznań bez przywołania żadnych kryteriów w przypadku zeznań uznanych za wiarygodne (6,4%) niż niewiarygodne (0,7%). W ocenie zeznań uznanych za wiarygodne częściej dochodziło do jej uzasadnienia z przewagą kryteriów przedmiotowych nad podmiotowymi w porównaniu z zeznaniami uznawanymi za niewiarygodne. Wreszcie, w przypadku zeznań niewiarygodnych istotnie częściej ocenę tę uzasadniano taką samą liczbą kryteriów przedmiotowych co podmiotowych (19,4%) w porównaniu z zeznaniami wiarygodnymi (4,1%, $\chi^2(3) = 42,06$, $p < 0,05$).

Następnie dokonano pogłębionej analizy dotyczącej średniej liczby kryteriów podmiotowych i przedmiotowych przywoływanych w uzasadnieniach wyroków sądowych.

jako sąd zlokalizowany w mieście będącym jednocześnie siedzibą sądu okręgowego) oraz małym sądzie rejonowym (SR w Brzesku) zlokalizowanym w mieście, które nie jest jednocześnie siedzibą sądu okręgowego.

Analiza wskazuje na statystycznie istotne różnice średniej liczby użytych kryteriów (zarówno przedmiotowych, jak i podmiotowych) w zależności od wielkości sądu, który dokonuje oceny ($Z = -4,19$, $p < 0,05$ oraz $Z = -5,03$, $p < 0,05$, odpowiednio). Innymi słowy – sędziowie orzekający w dużym sądzie (sądzie rejonowym w mieście będącym jednocześnie siedzibą sądu okręgowego) używają więcej kryteriów przedmiotowych i mniej podmiotowych niż sędziowie orzekający w małym sądzie (zlokalizowanym w mieście niebędącym siedzibą sądu okręgowego). Ponadto stwierdzono istotne statystycznie różnice średniej liczby przywołanych kryteriów przedmiotowych i podmiotowych w zależności od powagi przestępstwa będącego przedmiotem postępowania ($H = 14,34$, $p < 0,05$ i $H = 11,98$, $p < 0,05$, odpowiednio). Sądy przywołują istotnie mniej kryteriów przedmiotowych w ocenie zeznań złożonych w sprawach ciężkich niż lekkich występzków ($Z = -3,88$, $p < 0,05$). Różnica średniej liczby kryteriów przedmiotowych przywołanych w ocenie zeznań w sprawach ciężkich i średnich była bliska istotności statystycznej ($Z = -1,91$, $p = 0,56$). Sądy przywołują również istotnie mniej kryteriów przedmiotowych w ocenie zeznań złożonych w sprawach średnich niż lekkich występzków ($Z = -2,12$, $p < 0,05$). Stwierdzono ponadto, że sądy przywołują istotnie mniej kryteriów podmiotowych w sprawach ciężkich niż średnich i lekkich występzków ($Z = -3,38$, $p < 0,05$ oraz $Z = -2,61$, $p < 0,05$, odpowiednio). Różnica w liczbie zastosowanych kryteriów podmiotowych w sprawach średnich i lekkich występzków nie była statystycznie istotna.

Stwierdzono również istotne różnice w średniej liczbie kryteriów przedmiotowych i podmiotowych w zależności od charakterystyk przesłuchiwanego świadka. Sędziowie przywołują istotnie więcej kryteriów przedmiotowych w ocenie zeznań pokrzywdzonych niż świadków postronnych ($Z = -7,87$, $p < 0,05$). Ponadto wyniki sugerują, że sędziowie przywołują średnio więcej kryteriów przedmiotowych w ocenie zeznań świadków powiązanych z oskarżonym niż z nim niepowiązanych ($Z = -6,11$, $p < 0,05$). Analogiczną tendencję zaobserwowano w odniesieniu do kryteriów podmiotowych ($Z = -2,88$, $p < 0,05$).

Wreszcie, odnotowano, że sądy przywołują istotnie więcej kryteriów zarówno przedmiotowych, jak i podmiotowych, gdy uznają zeznania za niewiarygodne (lub częściowo niewiarygodne) niż gdy oceniają je jako wiarygodne ($Z = -2,51$, $p < 0,05$ oraz $Z = -7,46$, $p < 0,05$).

Czynniki wpływające na decyzję o wiarygodności lub braku wiarygodności zeznań świadków

Poszukiwanie odpowiedzi na drugie z pytań badawczych, dotyczące wpływu wybranych czynników na finalną ocenę zeznań świadka, ujawniło interesujące

rezultaty. Sędziowie rzadziej uznają zeznania pokrzywdzonych za niewiarygodne (15,1%) niż zeznania postronnych świadków przestępstwa (22,4%, $\chi^2(1) = 3,69$, $p < 0,05$).

Ponadto wszystkie ($N = 2$) zeznania małoletnich były uznane za niewiarygodne (lub częściowo niewiarygodne), podczas gdy zeznania osób starszych (65+) – w 21,1% przypadków i w 20,7% – zeznania osób dorosłych (16–64). Zaobserwowana zależność jest istotna statystycznie ($\chi^2(2) = 7,58$, $p < 0,05$). Ze względu jednak na to, że założenie testu χ^2 zostało złamane (33,3% komórek tabeli krzyżowej ma liczebność oczekiwaną mniejszą niż 5), wynik należy traktować z ostrożnością.

Wreszcie, zeznania osób powiązanych rodzinnie z oskarżonym były istotnie częściej uznawane za niewiarygodne (37,7%) niż zeznania świadków niepowiązanych z oskarżonym (17,4%, $\chi^2(1) = 22,17$, $p < 0,05$). Analogiczną tendencję zaobserwowano w ocenie zeznań świadków powołanych z inicjatywy oskarżonego lub jego pełnomocnika (30,6% zeznań uznanych za niewiarygodne w porównaniu z zeznaniami świadków powołanych z inicjatywy innych podmiotów [19,9%], choć różnica nie okazała się statystycznie istotna).

Dyskusja wyników

Średnia liczba świadków przesłuchanych w obu sądach rejonowych wskazuje, że zeznania świadków pozostają jednym z najistotniejszych środków dowodowych wykorzystywanych w sprawach karnych. Wydaje się, że to potwierdza powszechne w literaturze przekonanie o doniosłości tego środka dowodowego (np. de Poot, Bokhorst, van Koppen, Muller, 2004; Waltoś, 1985). Badania potwierdzają, że sędziowie i ławnicy (w systemie *common law*) polegają na dowodzie z zeznań świadków (np. Semmler, Brewer, Douglas, 2011; Vallas, 2011). Poza przedmiotem niniejszych badań było ustalenie struktury dowodów w analizowanych sprawach, nie sposób zatem przesądzić, czy zeznania świadków dominują nad innymi środkami dowodowymi przeprowadzonymi w sprawie i czy w ocenie dowodów przydawana jest im większa waga niż innym rodzajom dowodów.

Zastanawiające jest, że sądy przesłuchują tak znaczącą liczbę świadków, a jednocześnie ocena ich zeznań wydaje się powierzchowna. Po pierwsze, ponad 10% zeznań złożonych na rozprawie w ogóle nie została poddana ocenie przez sąd w uzasadnieniu wyroku. Po drugie, średnią liczbę kryteriów oceny zeznań przywołanych w uzasadnieniach trudno uznać za znaczącą. Sędziowie używali stosunkowo niewielu kryteriów przedmiotowych i jeszcze mniej podmiotowych. Zwłaszcza w kwestii kryteriów podmiotowych zaobserwować można szczególną niechęć sędziów do ich przywoływania w uzasadnieniach. Rezultat ten jest dość zaskakujący w świetle wyrażanego

wielokrotnie w orzecznictwie polskich sądów stanowiska o wadze, jaką dla oceny zeznań ma uwzględnienie zarówno treści zeznania, jak i osoby zeznającego świadka (wyrok SN z dnia 24.11.1975, II KR 254/75; wyrok SN z dnia 10.08.2005, II KK 410/04). Dlaczego więc sądy nie są skłonne powoływać się na kryteria podmiotowe w ocenie zeznań świadków? Być może sędziowie obawiają się uzasadniać swoje oceny argumentami, które mogą się wydawać „ulotne” i „nienamacalne”. Stosunkowo łatwo jest wykazać, że zeznanie jest wewnętrznie spójne (lub niespójne) albo zgodne (lub niezgodne) z pozostałym materiałem dowodowym poprzez wskazanie właściwych fragmentów protokołu przesłuchania albo innych dowodów. Trudniej natomiast wykazać, że świadek był (lub nie) pewny siebie czy też zeznawał spontanicznie (lub sprawiał wrażenie wyuczonego). Mimo to sędziowie nie powinni rezygnować z odwoływania się do kryteriów podmiotowych w ocenie zeznań. Nie pomijają podmiotowych aspektów (np. kwestii motywacji do złożenia zeznań) także psychologiczne narzędzia do oceny psychologicznej prawdziwości zeznań, jak choćby „Ocena prawdziwości zeznania” (*Statement Validity Analysis, SVA*; Memon, Vrij, Bull, 2003), tym samym nie powinny one być lekceważone w prawnej ocenie wiarygodności. Niemniej jednak nie ulega wątpliwości, że kryteria podmiotowe powinny stanowić uzupełnienie oceny treści zeznania na podstawie kryteriów o charakterze przedmiotowym.

Wreszcie, o powierzchowności prawnej oceny zeznań świadczy częstotliwość powoływania się na konkretne kryteria oceny zeznań. W uzasadnieniach dominowały trzy kryteria oceny o charakterze przedmiotowym – zgodność z pozostałym materiałem dowodowym, logiczność zeznań i ich wewnętrzna spójność, natomiast w ramach kryteriów podmiotowych – te odwołujące się do motywacji zeznającego świadka. Nie ulega wątpliwości, że największą wagę sądy przywiązują do kryterium zgodności z pozostałym materiałem dowodowym. Skłania je pewnie zresztą ku temu stanowisko Sądu Najwyższego, który w jednym z orzeczeń wskazał, że o wiarygodności zeznań świadka decyduje przede wszystkim to, czy znalazły one przekonujące potwierdzenie w pozostałym do dyspozycji sądu materiale dowodowym. Jedyne przy braku takiego potwierdzenia przy ocenie zeznań mogą mieć znaczenie także inne okoliczności (wyrok SN z dnia 28 maja 1979 roku, I KR 91/79). Wydaje się jednak, że takie stanowisko jest nadmiernie uproszczone i nie uwzględnia złożoności samych zeznań, jak i procesu ich oceny. Takie przekonanie – jeśli w istocie jest podzielane przez sędziów – może przyczyniać się niejako do lekceważenia wagi procesu oceny zeznań. Skoro dopiero stwierdzenie niezgodności zeznań z pozostałymi dowodami ma uruchamiać bardziej wnikliwą ocenę zeznań, to póki takiej oczywistej sprzeczności nie dostrzeżono, póty ocena pozostaje powierzchowna i ogólnikowa. To

niepokojące ze względu na wagę, jaką ocena dowodów odgrywa w procesie dokonywania ustaleń faktycznych w sprawach karnych.

Wciąż bez odpowiedzi pozostaje pytanie: czy faktyczna ocena zeznań jest tak powierzchowna, jak wynika z przeprowadzonej analizy uzasadnień wyroków sądowych? Można argumentować, że tak jest w istocie – ocena jest ogólnikowa, co znajduje odzwierciedlenie w uzasadnieniu, w którym sędziowie przywołują niewiele utartych kryteriów oceny. Można również postawić tezę, że ocena zeznań jest wnikliwa na etapie przesłuchania świadka czy przed podjęciem decyzji co do wyroku kończącego postępowanie, a jedynie pisemne uzasadnienie wyroku (które zwykle powstaje w kilka tygodni po ogłoszeniu wyroku) jest sztapkowe i ogólnikowe. Nie można wykluczyć, że pisemne uzasadnienie nie odzwierciedla (choć powinno) faktycznego przebiegu procesu oceny zeznań. Ta rozbieżność może wynikać z kilku przyczyn. Po pierwsze, z rutyny. Sędziowie przygotowują uzasadnienia w pewien schematyczny, sztapkowy sposób powielany w kolejnych uzasadnieniach. Po drugie, może być ona rezultatem *confirmation bias* (Nisbett, Ross, 1980; Rosenthal, 1994; Synder, Swan, 1978). Jeśli sąd na etapie wyrokowania podejmuje decyzję o winie (lub niewinności) oskarżonego, to przygotowując uzasadnienie, może mieć tendencję do interpretowania dowodów podlegających ocenie zgodnie z przekonaniem o (braku) odpowiedzialności za przestępstwo. Innymi słowy, jeśli sąd wydał wyrok skazujący, to może obdarzyć walorem wiarygodności te dowody (lub te aspekty dowodu np. zeznań świadka), które przemawiają za winą oskarżonego, a pomijać te środki dowodowe (lub ich aspekty), które są sprzeczne z tezą o odpowiedzialności oskarżonego za przestępstwo. Po trzecie wreszcie, powierzchowność oceny może wynikać z tego, że sędziowie kierują się intuicją w ocenie zeznań świadków, a jedynie *post factum* – przygotowując uzasadnienie – uzasadniają ocenę wiarygodności rozpoznawalnymi kryteriami oceny. Bez względu na rzeczywiste przyczyny wspomnianej ogólnikowości oceny wiarygodności zeznań świadków w uzasadnieniu pisemnym wyroku sądu rodzi ona poważne konsekwencje. Jeśli któraś ze stron postępowania decyduje się na złożenie apelacji, to w środku odwoławczym powinna wskazać zarzuty wobec zaskarżanego orzeczenia. Jeśli, dla przykładu, strona chce podnieść w apelacji, że sąd nieprawidłowo ocenił dowody (przydał walor wiarygodności zeznaniom świadka, które były w opinii strony niewiarygodne), to powinien móc wskazać błędy w rozumowaniu i ocenie sądu. Dokładna ocena, odzwierciedlona w uzasadnieniu pisemnym, umożliwi również sądowi wyższej instancji prześledzenie procesu oceny przeprowadzonej przez sąd niższej instancji. W postępowaniu apelacyjnym sądy rzadko prowadzą postępowanie dowodowe, a zwłaszcza w pełnym zakresie, w szczegól-

ności gdy dowód został już przeprowadzony przez sąd niższej instancji.

Wydaje się, że sędziowie w przypadku, gdy ostatecznie uznają zeznania za niewiarygodne (lub częściowo niewiarygodne), podchodzą do uzasadnienia procesu oceny zeznań nieco bardziej wnikliwie. Można odnieść zresztą wrażenie, że uznanie zeznań za wiarygodne jest pewnym standardem w polskim procesie karnym, skoro blisko 80% zeznań wszystkich świadków, których relacje zostały poddane ocenie, zostało uznane za wiarygodne. Kiedy natomiast sądy odmawiają waloru wiarygodności zeznaniom, wówczas podają więcej prawnych kryteriów oceny tego stanu rzeczy niż gdy uznają je za wiarygodne. Ponadto w tym przypadku sądy istotnie rzadziej (i generalnie niezwykle rzadko) w porównaniu z zeznaniami ocenionymi jako wiarygodne dopuszczają sytuację, by ocena zeznań nie pociągała za sobą żadnych kryteriów oceny. Sądy nie tylko podają w sytuacji oceny zeznań niewiarygodnych więcej kryteriów przedmiotowych, ale i podmiotowych, choć, co należy podkreślić, nie było statystycznie istotnej różnicy w liczbie zeznań, które zostałyby ocenione z przewagą kryteriów podmiotowych nad przedmiotowymi między zeznaniami uznanymi za wiarygodne i niewiarygodne. Co ciekawe wszakże, to fakt, że taka istotna i znacząca różnica dotyczyła zeznań, w których do oceny użyto takiej samej liczby kryteriów obu typów – do takiej sytuacji dochodziło częściej w przypadku zeznań uznanych za niewiarygodne. Wydaje się, że to potwierdza wyrażone przez Płachtę (1985) przekonanie, że sądy mają tendencję do uzasadniania braku wiarygodności zeznań kryteriami o charakterze podmiotowym, a kryteriami przedmiotowymi – ich wiarygodności. Zdaniem Płachty (1985) wynika to z faktu, że gdy sędziowie uznają zeznania za wiarygodne, sięgają do treści dowodu, odnajdując w nim wystarczające podstawy do swojej oceny. Natomiast gdy poddają w wątpliwość wiarygodność zeznania, sięgają po dodatkowe argumenty podważające analizowany środek dowodowy – w tym przypadku niejako „zewnątrzny” w stosunku do treści dowodu, dotyczący jego źródła.

Analiza uzyskanych wyników prowadzi do interesujących wniosków co do wpływu pewnych charakterystyk zarówno spraw karnych, jak i samych świadków na ocenę zeznań. Sędziowie używają istotnie mniej kryteriów oceny zeznań złożonych w sprawach o poważne występki w porównaniu z występkami średnimi i lekkimi. Co więcej, dotyczy to zarówno liczby przedmiotowych, jak i podmiotowych kryteriów oceny zeznań złożonych w postępowaniach dotyczących występków poważnych. To może niepokoić szczególnie, gdyż w sprawach o takie przestępstwa, jak nadużycia seksualne wobec małoletnich, oszustwa czy rozboje, oczekuje się od sądów orzekających w tych kategoriach spraw szczególnie skrupulatnej oceny dowodów. Pozostaje mieć nadzieję (choć relacjonowane badania nie dostarczają dowodów na po-

parcie tej tezy), że im poważniejsze przestępstwo, w tym większym stopniu sądy polegają na innych niż osobowe środkach dowodowych. Być może w przypadku przestępstw poważnych także działania organów ścigania na etapie postępowania przygotowawczego są ukierunkowane z większą intensywnością na poszukiwanie dowodów rzeczowych, które potem na rozprawie dominują w postępowaniu dowodowym.

W odniesieniu do rodzajów kryteriów przywoływanych przez sądy zaobserwowano różnice w częstotliwości powoływania się na rodzaj kryteriów oceny w zależności od wielkości sądu. W Sądzie Rejonowym dla Krakowa-Nowej Huty (tzw. dużym sądzie z racji lokalizacji w mieście, będącym jednocześnie siedzibą sądu okręgowego) istotnie częściej posługiwano się w ocenie zeznań kryteriami o charakterze przedmiotowym, a rzadziej podmiotowym, niż w Sądzie Rejonowym w Brzesku (tzw. małym sądzie, czyli będącym wyłącznie siedzibą sądu tego szczebla). W zeznaniach ocenianych przez sędziów w „małym” sądzie rejonowym istotnie częściej zdarzały się takie zeznania, w ocenie których przeważały kryteria podmiotowe nad przedmiotowymi oraz takie, w których liczba kryteriów obu typów była równa, niż w „dużym” sądzie. Trudno jednoznacznie przesądzić, dlaczego w sądzie w dużym ośrodku zaobserwowany został większy dystans wobec kryteriów podmiotowych w porównaniu z małym ośrodkiem.

Okazuje się, że rola świadka w postępowaniu ma istotny wpływ na liczbę oraz rodzaj powoływanych kryteriów oceny zeznań. Po pierwsze, sędziowie powołują więcej kryteriów oceny, gdy oceniają zeznania pokrzywdzonego, niż w przypadku oceniania zeznań postronnego świadka. Ponadto w ocenie zeznań pokrzywdzonych częściej posługują się kryteriami przedmiotowymi niż podmiotowymi w porównaniu z zeznaniami świadków postronnych. Wreszcie – wyniki sugerują, że zeznania pokrzywdzonych są częściej uznawane za wiarygodne niż zeznania świadków postronnych. Innymi słowy – sądy dokonują bardziej wnikliwej oceny zeznań pokrzywdzonych, są bardziej skłonne uzasadniać zeznania pokrzywdzonych, powołując się na kryteria przedmiotowe (a zatem odnoszące się do treści zeznań), i co więcej – częściej niż świadków postronnych uważają ich za wiarygodnych. Wydaje się zatem, że sędziowie mają większe zaufanie do zeznań pokrzywdzonych niż świadków postronnych, ale wnikliwiej oceniają, czy zeznania ofiar przestępstw rzeczywiście na to zaufanie zasługują. Ponadto ocenę tę uzasadniają bardziej „uchwytnymi” i „namacalnymi” argumentami o charakterze przedmiotowym raczej niż argumentami osobowymi. Być może sędziowie mają przekonanie, że motywacja pokrzywdzonych do złożenia zeznań, w tym zeznań określonej treści, jest w większości przypadków podyktowana (uważanym za wątpliwe) zainteresowaniem określonym rezultatem postępowania (najczęściej skazaniem oskarżonego).

To w odczuciu sędziów może, niejako automatycznie, rzucić pewien cień na wiarygodność pokrzywdzonego. Być może z tego względu właśnie sędziowie są bardziej skłonni uzasadniać wiarygodność zeznań pokrzywdzonych, powołując się na kryteria, w ich ocenie, bardziej przekonujące.

Pewne ostrożne wnioski można wysnuć w odniesieniu do wpływu wieku na ocenę zeznań świadków w świetle prezentowanych wyników. Po pierwsze, odnotowano, że im starszy świadek, tym mniej kryteriów oceny zeznań powołują sędziowie. Związek ten okazał się słaby, ale sugeruje, że sądy dokonują bardziej wnikliwej oceny zeznań młodszych świadków, w tym dzieci. Liczne badania nad zeznaniami dzieci (przegląd badań można znaleźć w pracy autorstwa Principe, Greenhoot i Ceci, 2014) wskazują, że dzieci, w szczególności młodsze, dostarczają mniej informacji niż dorośli i dzieci starsze (Pipe, Lamb, Orbach, Esplin, 2004), częściej nieprawidłowo relacjonują wiek, wagę i wzrost sprawcy przestępstwa (Pozzulo, 2007) i są bardziej podatne na sugestie niż dorośli (Melnik, Crossman, Scullin, 2007). Być może sędziowie – świadomi tych ograniczeń pamięci i zeznań dzieci w roli świadków – bardziej wnikliwie podchodzą do zeznań małoletnich. Niepokoi wszakże, że podobna uważność nie towarzyszy ocenie zeznań świadków w podeszłym wieku, których sposób poznawczego funkcjonowania odbiega od funkcjonowania osób dorosłych (Mueller-Johnson, Ceci, 2007). Zainteresowanie badaczy pamięcią i zeznaniami osób starszych jest jednak zdecydowanie „świeższe” niż zeznaniami dzieci i, być może, wiedza ta nie przeniknęła jeszcze do systemu wymiaru sprawiedliwości. Nieco bardziej niepokojące są obserwacje dotyczące wpływu wieku na ocenę zeznań jako wiarygodnych lub niewiarygodnych. Okazuje się, że zeznania świadków do 15. roku życia były częściej niż u pozostałych kategorii świadków uznawane za niewiarygodne. Ze względu jednak na niespełnienie wszystkich wymogów analizy statystycznej należy wyciągać bardzo ostrożne wnioski. Zawodowi uczestnicy procesu karnego niejednokrotnie wskazują, że zeznania dzieci są mniej godne zaufania niż dorosłych (Benton, Ross, Bradshaw, Thomas, Bradshaw, 2006; Granhag i in., 2005). W wielu jurysdykcjach pokutowała przez wiele lat – i być może wciąż pokutuje – podejrzliwość wobec zeznań tej kategorii świadków (Bala, Ramakrishnan, Lindsay, Lee, 2005). Nie można wykluczyć, że ta podejrzliwość przenika również polski system wymiaru sprawiedliwości karnej.

Zaobserwowano również, że sądy istotnie więcej kryteriów oceny zeznań przywołują wówczas, gdy oceniają zeznania świadków przesłuchanych z udziałem biegłego psychologa niż bez jego udziału. Zgodnie z art. 192 § 2 k.p.k. jeżeli istnieje wątpliwość co do stanu psychicznego świadka, jego stanu rozwoju umysłowego, zdolności postrzegania lub odtwarzania przez niego postrzeżeń, sąd lub prokurator może zarządzić przesłu-

chaniu świadka z udziałem biegłego lekarza lub biegłego psychologa, a świadek nie może się temu sprzeciwić. Art. 194 § 4 k.p.k. wskazuje ponadto, że świadka można poddać badaniu psychologicznemu. W praktyce zatem przesłuchanie z udziałem biegłego obejmuje również sporządzenie opinii sądowo-psychologicznej dotyczącej psychologicznej prawdziwości złożonych zeznań. Większa liczba przywołanych kryteriów oceny zeznań w przypadku świadków, których zeznania poddane zostały również ocenie psychologa, prawdopodobnie wynika z tego, że sędziowie czerpią z opinii psychologicznych argumenty na poparcie swojej – prawnej – oceny dowodu z zeznań świadka. Wiele kryteriów prawnej oceny zbieżnych jest zresztą z kryteriami psychologicznej prawdziwości zeznań wskazanymi w najpowszechniej stosowanym narzędziu do jej oceny, a mianowicie wspomnianej już „Ocenie prawdziwości zeznania” (*Statement Validity Analysis, SVA*). SVA składa się z kilku elementów, spośród których najistotniejsza jest oparta na kryteriach analiza treści (*Criteria-Based Content Analysis, CBCA*), która zawiera 19 kryteriów treściowych prawdziwości zeznania. Dla przykładu: kryterium treściowym występującym w CBCA jest „logiczna struktura” (*logic structure*), jest ono spełnione, gdy relacja zasadniczo ma sens, jest spójna i logiczna, a poszczególne jej części nie są ze sobą sprzeczne (Memon i in., 2003), odpowiada prawnym kryteriom logiczności zeznań i wewnętrznej spójności. Z kolei kryteria CBCA – „duża liczba szczegółów” (*appropriate quantity of detail*) oraz „nieustrukturyzowany sposób prezentacji” (*unstructured production*) – odpowiadają prawnym kryteriom szczegółowości zeznań i spontaniczności zeznań.

Ciekawe rezultaty zaobserwowano w odniesieniu do oceny zeznań świadków powiązanych rodzinnie z oskarżonym. Sędziowie w uzasadnieniach podają więcej kryteriów, kiedy oceniają zeznania świadków będących osobami najbliższymi dla oskarżonego (w rozumieniu osób uprawnionych do odmowy składania zeznań na podstawie art. 182 k.p.k.) w porównaniu ze świadkami, których nie łączy z oskarżonym ten rodzaj relacji. Ponadto, w przypadku zeznań osób najbliższych oskarżonemu istotnie rzadziej, w porównaniu ze świadkami z nim niepowiązanymi, zdarzało się, by oceny zeznań sąd nie uzasadnił żadnymi jej kryteriami. Sądy w odniesieniu do tej kategorii świadków powoływały zarówno więcej kryteriów o charakterze przedmiotowym, jak i podmiotowym. Można domniemywać, że oceniając zeznania osób najbliższych oskarżonemu, sędziowie dokonują bardziej wnikliwej analizy ich zeznań. Co ciekawe – bez względu na wnikliwość tej oceny zeznania osób najbliższych oskarżonemu były istotnie częściej uznawane za niewiarygodne niż zeznania osób z nim niezwiązanych. To może sugerować pewne uprzedzenie sądów do świadków będących osobami najbliższymi oskarżonemu, choć nie ulega wątpliwości, że powinny one podchodzić do

każdego rodzaju dowodu obiektywnie i bez uprzedzeń (Płachta, 1985). Naiwnością byłoby jednak sądzić, że sędziowie nie ulegają różnego rodzaju uprzedzeniom i że nie wążą one na podejmowanych w procesie karnym decyzjach, zwłaszcza gdy dotyczą one osoby oskarżonego (Rachlinski, Johnson, Wistrich, Guthrie, 2009; Wistrich, Rachlinski, 2017). Niemniej jednak, skoro możliwa jest stronniczość sądów ze względu na pewne charakterystyki czy właściwości oskarżonego, to zapewne możliwe jest, że ulegają one uprzedzeniom względem świadków (a tym samym – pośrednio – uprzedzenia te mogą wpływać na ocenę odpowiedzialności oskarżonego). Trzeba oczywiście mieć na względzie, że w przypadku sędziów wiele ze środków, które w literaturze uchodzą za takie, które minimalizują ich wpływ na podejmowane decyzje, znajduje już zastosowanie – sędziowie są m.in. dobrze wykształceni i wyszkoleni, mają doświadczenie w podejmowaniu trudnych decyzji, są umotywowani, aby ich decyzje były słuszne. Ponoszą odpowiedzialność za swoje decyzje: podejmują je publicznie, ich decyzje podlegają kontroli apelacyjnej, a ich awans zawodowy jest poprzedzony analizą dotychczasowego orzecznictwa (Wistrich, Rachlinski, 2017).

Doniesienia z zakresu psychologii i prawa wskazują, że dowód z zeznań świadków (a przede wszystkim błędy w zeznaniach, a także zeznania fałszywe) stanowi jedną z kluczowych przyczyn, dla których systemy prawne na całym świecie borykają się z problemem pomyłek sądowych (np. Huff, 1987; Rattner, 1988). Powierzchność oceny zeznań przez sądy, zidentyfikowana w relacjonowanych badaniach empirycznych, może do pewnego stopnia tłumaczyć, dlaczego tak często do niesłusznych skazań dochodzi właśnie w rezultacie nieprawdziwych lub fałszywych zeznań (np. Scheck, Neufeld, Dwyer, 2000). Powinno to jednak szczególnie uwrażliwić sędziów na znaczenie wnikliwej i szczegółowej oceny zeznań. Fakt, że ocena zeznań (i innych dowodów w postępowaniu) ma charakter prawny, nie oznacza, że może się ona odbywać w oderwaniu od współczesnej wiedzy psychologicznej. Nie sposób wyeliminować błędów popełnianych przez naocznych świadków przestępstw, te zawsze będą się zdarzać. Niemniej można ograniczyć ich wpływ na postępowanie karne, do czego wiedza sędziów umiejętnie stosowana w praktyce może się istotnie przyczynić.

Podziękowania

Badania zostały sfinansowane przez Narodowe Centrum Nauki (grant numer 2015/19/D/HS5/01241).