*German Medical Data Sciences 2022 - Future Medicine*
*R. Röhrig et al. (Eds.)*


*doi:10.3233/SHTI220798*

# The Interpretation of Verbal Probabilities: A Systematic Literature Review and Meta-Analysis

Hannah VOGEL[a], Sebastian APPELBAUM[a,b], Heidemarie HALLER[c], Thomas OSTERMANN[a,1]

[a] *Department for Psychology and Psychotherapy, Witten/Herdecke University, Alfred Herrhausen-Straße 50, 58448 Witten, Germany*

[b] *Trimberg Research Academy, University of Bamberg, An der Weberei 5, 96047 Bamberg, Germany*

[c] *Department of Internal and Integrative Medicine, Evang. Kliniken Essen-Mitte, Faculty of Medicine, University of Duisburg-Essen, Am Deimelsberg 34a, 45276 Essen, Germany*

**Abstract. Introduction:** Verbal probabilities such as "likely" or "probable" are commonly used to describe situations of uncertainty or risk and are easy and natural to most people. Numerous studies are devoted to the translation of verbal probability expressions to numerical probabilities. **Methods:** The present work aims to summarize existing research on the numerical interpretation of verbal probabilities. This was accomplished by means of a systematic literature review and meta-analysis conducted alongside the MOOSE-guidelines for meta-analysis of observational studies in epidemiology. Studies were included, if they provided empirical assignments of verbal probabilities to numerical values. **Results:** The literature search identified 181 publications and finally led to 21 included articles and the procession of 35 verbal probability expressions. Sample size of the studies ranged from 11 to 683 participants and covered a period of half a century from 1967 to 2018. In half of the studies, verbal probabilities were delivered in a neutral context followed by a medical context. Mean values of the verbal probabilities range from 7.24% for the term "impossible" up to 94.79% for the term "definite". **Discussion:** According to the results, there is a common 'across-study' consensus of 35 probability expressions for describing different degrees of probability, whose numerical interpretation follows a linear course. However, heterogeneity of studies was considerably high and should be considered as a limiting factor.

**Keywords.** Verbal probabilities, Meta-analysis, Numerical representation, Systematic review

## 1. Introduction

Within the last decades, the use and interpretation of verbal probability expressions (VPE) has been intensively investigated from different perspectives such as the field of economics, politics or the health sector [1-5]. VPE are commonly used to describe situations of uncertainty or risk and according to [6] are easy and natural to most people.

[1] Thomas Ostermann,Department of Psychology and Psychotherapy, Witten/Herdecke University, Alfred-Herrhausen-Straße 50, 58448 Witten, Germany; E-mail: thomas.ostermann@uni-wh.de.

Brun and Teigen (1988) observed that physicians preferred communicating probabilities verbally whereas their patients rather preferred receiving health-related information numerically [2]. A problem in the transfer of verbal into numerical probabilities is the considerable between-subject variability in the use and interpretations of VPE [1,7]. For example, wide between-subject variations were found for the expression "likely", which was assigned probability estimates ranging from p = .5 to p = .95 [8].

Concurrently, individuals show an internal consistence in the use and interpretation of VPE [9,10]. Dhami and Wallsten (2005) linked those findings and argued, that individuals have a stable lexicon of VPE, which however may differ considerably among individuals [7]. Still, interpersonal consensus regarding the probabilistic meaning of an expression may be derived by examining the rank order of the expression within each lexicon. The authors suggest that expressions or phrases, which are ranked equally, are likely to have similar meanings even though the exact wording differs among individuals. These results are supported in [11], where the authors observed the rank order of a set of 23 VPE to be relatively stable in a British, Hellenic and Malaysian sample. However, the numerical values that were assigned to each VPE turned out to differ considerably between nationalities.

A central concern in the interpretational variability of VPE thus is the danger of communicative misunderstandings [12]. According to [2] most individuals are unaware of both the ambiguity of VPE as well as the variability of VPE interpretations in the general population. Hence, various examinations have focused on providing a translation aid from verbal to numerical probabilities or vice versa [8,9,10,13,14,15].

The current work addresses this translational issue and aims at systematically reviewing and summarizing the existing literature about numerical interpretations of VPE in order to provide an overview of previous research results. In addition to the identification and description of studies involving the numerical interpretation of verbal probabilities, numerical interpretations of frequently examined VPE will be summarized statistically in order to generate a bundled numerical interpretation of each identified VPE.

## 2. Methods

This systematic review and meta-analysis was conducted alongside the guidelines for meta-analysis of observational studies in epidemiology (MOOSE) [16]. The following electronic databases were searched from their inception to 2020 independently by two authors: Psychological and Behavioral Science Collection (PBSC), PubMed, PsycArticles and CINAHL. The literature search was constructed around the search term 'verbal probabilities' and adapted for each database as necessary. Furthermore, reference lists of identified original articles and reviews were searched manually for further relevant articles.

Articles were included, if they studied empirical assignments of verbal probabilities to numerical values (from 0% to 100% or 0 to 1). Articles including numerical interpretation of verbal probabilities using another numerical format (e.g. five-point Likert scales or membership functions) were not included. Expressions such as 'rare', 'commonly' or 'often' referring to frequencies rather than probabilities and were not considered in the current work. Furthermore, only studies published in German, English, or Spanish were considered. Book chapters and unpublished studies were excluded.

Included studies were analyzed regarding the year of publication, country, method of capturing the probability interpretation, numerical format, and the thematic context of

the verbal probability assessment. Furthermore, the total amount of VPE examined in the studies and identified mean values, standard deviations as well as sample size for each study were extracted.

To review expressions that are used with a certain consensus, only VPE that were interpreted in at least four investigations were further processed. A meta-analysis for each VPE using the random effects model was calculated by R package 'meta' [17]. In cases of missing standard deviations, they were calculated by the average standard deviation of the remaining studies within the corresponding VPE. Missing sample sizes were substituted using the median of reported sample sizes of the remaining studies.

## 3. Results

The database search revealed 163 articles. Reviewing reference lists of identified articles led to the consideration of 18 further articles. After removal of duplicates, 87 articles were excluded by screening title and abstracts. Full texts of the remaining 49 records were reviewed and assessed for eligibility. Further articles had to be excluded as the respective full-text was not available, published in a different language, had a different numerical format, or did not present detailed results for meta-analysis. The selection process is illustrated in figure 1.

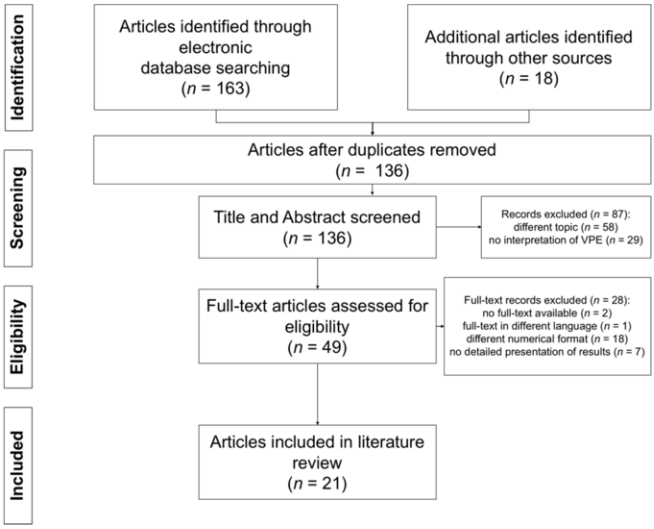


**Fig 1.** Flow chart of the literature selection process. n = here: number of articles; VPE = verbal probability expressions.

Table 1 provides an overview of the 21 included studies. Sample size of the studies ranged from 11 to 683 participants (Mean: 111.75 Median: 70) and covered a period of half a century from 1967 to 2018. Origin of the studies was mixed with n=10 studies from the USA followed by the UK (n=3) and Norway (n=2). Most of the studies (n=11) used a population of students or medical staff (n=4). Only one study asked a sample of patients. The VPE in half of the studies (n=10) were delivered in a neutral context followed by a medical context (n=7).

**Table 1.** Overview of the included studies.

| First Author | Year | Origin | N | Sample | Context |
|---|---|---|---|---|---|
| Bergenstrom [18] | 2003 | UK/USA | 87 | Medical students | Medical |
| Brun [2] | 1988 | Norway | 16 | Psychology students | Neutral |
| Budescu [9] | 1985 | USA | 32 | Psychology students | Neutral |
| Chee [11] | 2006 | Malaysia | 32 | Mixed (mainly students) | Neutral |
| Cohn [4] | 2009 | Mexico/USA | 263 | Mixed (mainly students) | Medical |
| Damrosch [19] | 1983 | USA | 70 | Female nurses | Medical |
| Hamm [13] | 1991 | USA | 140 | Psychology students | Neutral |
| Hobby [20] | 2000 | UK | 11 | Physicians | Medical |
| Honda [21] | 2006 | Japan | 137 | Students | Gambling |
| Honda [21] | 2006 | Japan | 67 | Students | Gambling |
| Juhanchich [22] | 2013 | USA | 84 | Workers | Neutral |
| Kong [12] | 1986 | USA | n/a | Medical staff | Medical |
| Lichtenstein [15] | 1967 | USA | 188 | Employees | Neutral |
| Ostermann [23] | 2018 | Germany | 683 | Mixed (mainly students) | Neutral |
| Reagan [8] | 1989 | USA | 115 | Psychology students | Neutral |
| Shying [24] | 2013 | HK, MYS, SGP | 55 | Auditors | Neutral |
| Sutherland [25] | 1991 | Canada | 100 | Cancer patients | Medical |
| Tavana [26] | 1997 | USA | 30 | Financial experts | Banking |
| Teigen [27] | 2001 | Norway | 20 | Psychology students | Job offer |
| Teixera [5] | 2009 | Portugal | 35 | Auditors | Neutral |
| Villejoubert [28] | 2009 | UK | 70 | Medical staff | Medical |

Statistical analyses were based on the following 35 VPE, which had been examined by at least four included articles: almost certain (n = 7), almost impossible (n = 5), certain (n = 7), chance (n = 4), definite (n = 4), doubtful (n = 4), good chance (n = 6), highly improbable (n = 6), highly probable (n = 8), impossible (n = 5), improbable (n = 8), likely (n = 18), maybe (n = 6), not certain (n = 7), not likely (n = 4), not possible (n = 5), not probable, (n = 4), perhaps (n = 4), possible (n = 22), possibly (n = 4), probable (n = 19), quite likely (n = 7), quite probable (n = 4), quite unlikely (n = 9), reasonable assurance (n = 4), reasonably certain (n = 4), reasonably possible (n = 4), remote (n = 4), somewhat doubtful (n = 4), uncertain (n = 11), unlikely (n = 14), very likely (n = 8), very probable (n = 10), very unlikely (n = 10), virtually certain (n = 4). Figures 2 to 4 provide forest plots for the VPE “unlikely”, “uncertain” and “likely”.

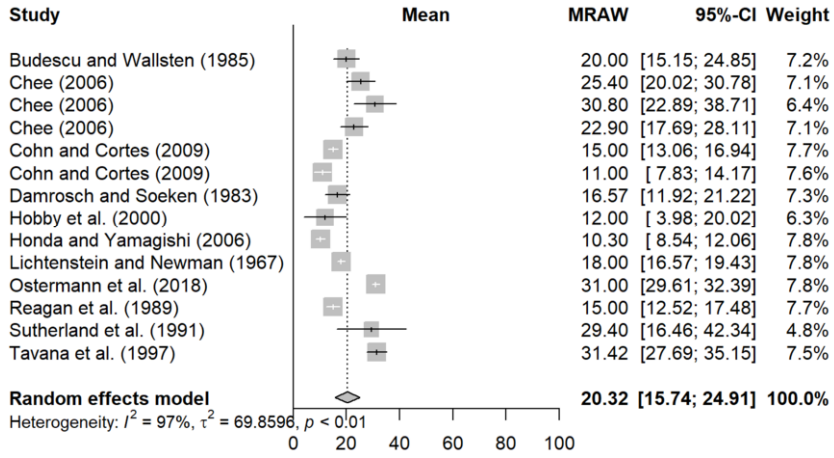


**Fig 2.** Forest plot of the meta-analysis of the expression “unlikely” from (MRAW: Raw Mean).

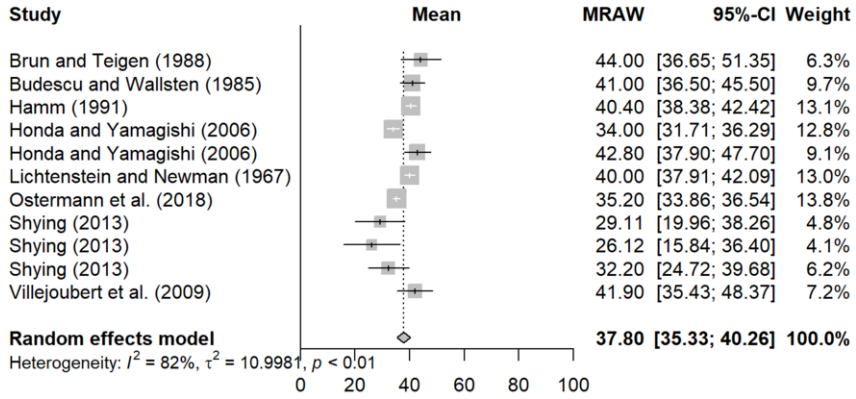


**Fig 3.** Forest plot of the meta-analysis of the expression "uncertain" (MRAW: Raw Mean).

Between-study heterogeneity for each was highly significant in 29 out of 35 cases except for the expressions "remote" ($I^2$= 0.00; $\tau^2$= 0.00; p = 0.604), "somewhat doubtful" ($I^2$;= 2.60; $\tau^2$= 0.68; p = 0.380), "reasonably possible" ($I^2$;= 15.40; $\tau^2$= 2.52; p = 0.315), "reasonable assurance" ($I^2$;= 40.70; $\tau^2$= 3.70; p = 0.168), "reasonably certain" ($I^2$;= 3.30; $\tau^2$= 0.21; p = 0.376) and "virtually certain" ($I^2$;= 64.80; $\tau^2$= 11.82; p = 0.036).

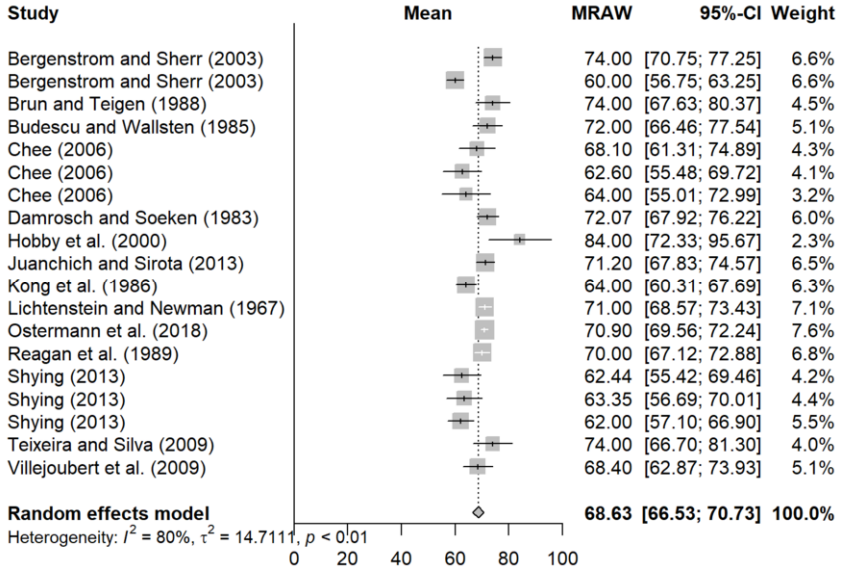


**Fig 4.** Forest plot of the meta-analysis of the expression "likely" (MRAW: Raw Mean).

The complete results of the random effects model for all VPEs are provided in Fig. 5, in which the weighted mean with 95% confidence interval is displayed.

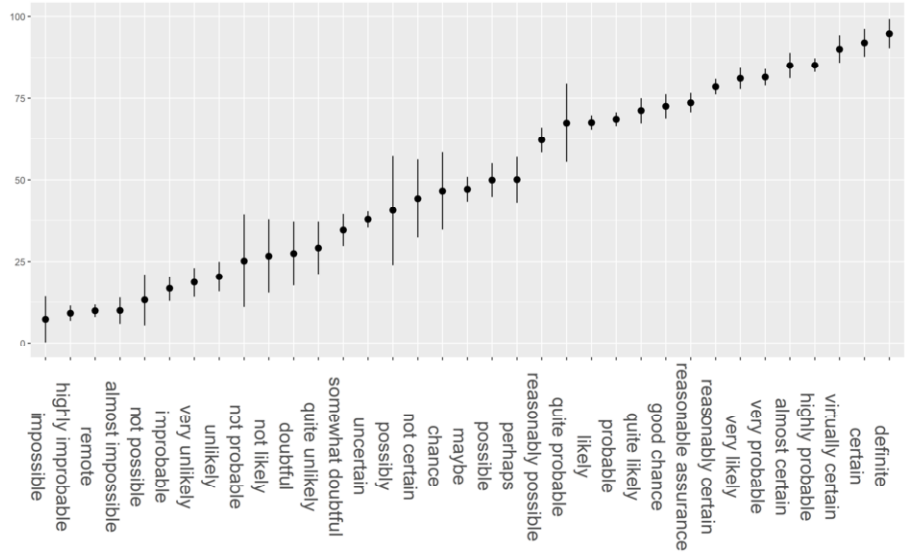


**Fig 5.** Averaged mean values of 35 verbal probability expressions across the respective studies; error bars denote the 95% confidence intervals.

Mean values of the VPE interpretations range from 7.24% (SE: 3.58) for the VPE "impossible" up to 94.79% (SE: 2.26) for the VPE "definite". The greatest leap can be found between the expressions "perhaps" (M: 50.05; SE: 3.57) and "reasonably possible" (M: 62.22; SE: 1.98). Nevertheless, the 35 expressions are apparently linear distributed across the continuum between 0 and 100%.

## 4. Discussion

This work aimed at systematically reviewing the existing literature on the interpretation of verbal probabilities. Studies including a numerical interpretation of VPE were included and further analyzed. A large number of expressions are compound words, consisting of a central expression of probability accompanied by a further expression used to describe the degree of probability more specifically such as barely possible, entirely possible or faintly possible. This results in innumerable differentiations of probability expressions compared to the central expression, for example 'possible'.

The extracted set of 35 VPE can thus be cautiously considered a common 'across-study' lexicon of verbal probability terms, and is evenly distributed within the range between 7.24% (impossible) up to 94.79% (definite). The present results provide strong evidence for VPE interpretations behaving in a linear rather than somehow logistic way.

Although this meta-analysis has its strength in synthesizing more than 40 years of research in numerical estimations of verbal probabilities, it also has several limitations. Firstly, we only searched for the term "verbal probabilities". Although this is the most common term in this field of research, we might have missed literature by using only a single term. We also decided against a risk of bias assessment, which normally attributes issues like "random sequence generation", "allocation concealment", "blinding" or

"selective reporting" which is an important issue in RCTs. Our analysis however only deals with surveys without control groups and blinding and thus we cannot give information on a risk of bias. Another major limitation is given by the heterogeneity of studies, indicated by means of $I^2$ and $\tau^2$, which is considerably high with 29 out of 35 comparisons reaching statistical significance. Including and comparing heterogeneous studies is one of the main criticisms regarding the method of a meta-analysis casually known as the 'apples and oranges problem' [29]. We however decided to include the studies in this investigation, as we just aimed at summing up existing numerical interpretations on VPE. The inclusion criteria set out above served the purpose to identify studies with comparable research questions and methodology.

A further limitation of the present meta-analysis is the combination of studies from different research contexts as well as uncontextualized investigations. Since previous examinations demonstrated the context of VPE to influence their numerical interpretation [1,2] future research should combine investigations using comparable contexts or address different contexts in subgroup analyses.

## 5. Conclusion

Our results systematically reviewed the existing literature on VPE and revealed, that a considerable amount of studies addresses the numerical interpretation of VPE; however, using different numerical formats. Considering studies using the format of percentages or probabilities leads to the identification of a broad variety of different probability expressions that were apparently linear distributed within the range from 0% to 100%. According to our results, there is a common 'across-study' consensus of 35 VPE for describing different degrees of probability, whose numerical interpretation follows a linear course as discussed in [23].

## 6. Conflict of Interest

The authors state that they have no conflict of interests.

## 7. Author's contribution

Conceptualization: Thomas Ostermann, Hannah Vogel, Heidemarie Haller; Methodology: Thomas Ostermann, Sebastian Appelbaum; Formal analysis: Sebastian Appelbaum; Writing – original draft: Hannah Vogel, Thomas Ostermann; Writing – review & editing: Thomas Ostermann, Heidemarie Haller, Sebastian Appelbaum, Hannah Vogel; Supervision: Thomas Ostermann.

## References

[1] R. Beyth-Marom. How probable is probable? A numerical translation of verbal probability expressions. *Journal of Forecasting*, (1982), **1(3)**, 257–269.

[2] W. Brun, K.H. Teigen. Verbal probabilities: ambiguous, context-dependent, or both? *Organizational Behavior and Human Decision Processes*, (1988), **41**, 390–404.

[3] D.V. Budescu, T.M. Karelitz, T.S. Wallsten. Predicting the directionality of probability words from their membership functions. *Journal of Behavioral Decision Making*, (2003), **16(3)**, 159–180.
[4] L. Cohn, M. Cortes. Quantifying risk: verbal probability expressions in Spanish and English. *American Journal of Health Behavior*, (2009), **33(3),** 244–255.
[5] C. Teixeira, A.F. Silva. The interpretation of verbal probability expressions used in the IAS/IFRS: some Portuguese evidence. *Polytechnical Studies Review*, (2009), **7(12)**, 57–73.
[6] A. C. Zimmer. A model for the interpretation of verbal predictions. *International Journal of Man-Machine Studies*, (1984), **20(1)**, 121–134.
[7] M. K. Dhami, T. S. Wallsten. Interpersonal comparison of subjective probabilities: toward translating linguistic probabilities. *Memory & Cognition*, (2005), **33(6)**, 1057–1068.
[8] R. T. Reagan, F. Mosteller, C. Youtz. Quantitative meanings of verbal probability expressions. *Journal of Applied Psychology*, (1989), **74(3)**, 433–442.
[9] D. V. Budescu, T. S. Wallsten. Consistency in interpretation of probabilistic phrases. *Organizational Behavior and Human Decision Processes*, (1985), **36(3)**, 391–405.
[10] D. A. Clark. Verbal uncertainty expressions: a critical review of two decades of research. *Current Psychology: Research & Reviews*, (1990), **9(3)**, 203–235.
[11] C. S. Chee. "In-Isolation" study on verbal uncertainty expressions. *Proceedings of the 2nd IMT-GT Regional Conference on Mathematics, Statistics and Applications,* (2006), University Sains Malaysia, Penang.
[12] A. Kong, G. O. Barnett, F. Mosteller, C. Youtz. How medical professionals evaluate expressions of probability. *The New England Journal of Medicine*, (1986), **315(12),** 740–744.
[13] R. M. Hamm. Selection of verbal probabilities: a solution for some problems of verbal probability expression. *Organizational Behavior and Human Decision Processes*, (1991), **48(2)**, 193–223.
[14] C. Witteman, S. Renooij. Evaluation of a verbal–numerical probability scale. International *Journal of Approximate Reasoning*, (2003), **33(2)**, 117–131.
[15] S. Lichtenstein, J. R. Newman. Empirical scaling of common verbal phrases associated with numerical probabilities. *Psychonomic Science*, (1967), **9(10)**, 563–564.
[16] D. F. Stroup, J. A. Berlin, S. C. Morton, I. Olkin, G. D. Williamson, D. Rennie, D. Moher, B. J. Becker, T. A. Sipe, S. B. Thacker. Meta-analysis of observational studies in epidemiology: a proposal for reporting. *JAMA*, (2000), **283(15)**, 2008-2012.
[17] S. Balduzzi, G. Rücker, G. Schwarzer. How to perform a meta-analysis with R: A practical tutorial. *Evidence-Based Mental Health*, (2019), **22(4)**, 153–160.
[18] A. Bergenstrom, L. Sherr. The effect of order of presentation of verbal probability expressions on numerical estimates in a medical context. *Psychology, Health & Medicine*, (2003), **8(4)**, 391–398.
[19] S. P. Damrosch, K. Soeken. Communicating probability in clinical reports: Nurses' numerical associations to verbal expressions. *Research in Nursing & Health*, (1983), **6(2)**, 85–87.
[20] J. L. Hobby, B. D. Tom, C. Todd, P.W. Bearcroft, A. K. Dixon. Communication of doubt and certainty in radiological reports. *The British Journal of Radiology*, (2000), **73(873)**, 999–1001.
[21] H. Honda, K. Yamagishi. Directional verbal probabilities: Inconsistencies between preferential judgments and numerical meanings. *Experimental Psychology*, (2006), **53(3)**, 161–170.
[22] M. Juanchich, M. Sirota. Do people really say it is "likely" when they believe it is only "possible"? Effect of politeness on risk communication. *Quarterly Journal of Experimental Psychology*, (2013), **66(7)**, 1268–1275.
[23] T. Ostermann, H. Vogel, S. Appelbaum. Verbal probabilities: linear or logistic? - A regression analysis approach. *Studies in Health Technology and Informatics*, (2018), **253**, 117–121.
[24] M. Shying. Auditors interpretations of “in-isolation” verbal probability expressions: A cross-national study. In M. Kennelley: The Influence of Environmental and Cultural Factors in International Accounting. Symposium conducted at the American Accounting Associations (AAA) Annual Meeting, Philadelphia. (2013). Available at: https://www.sec.gov/rules/concept/s70400/shying1.htm
[25] H. J. Sutherland, G. A. Lockwood, D. L. Tritchler, F. Sem, L. Brooks, J. E. Till. Communicating probablistic Information to cancer patients: is there 'noise' on the line? *Social Science & Medicine*, (1991), **32(6)**, 725–731.
[26] M. Tavana, D. Kennedy, B. Mohebbi. An Applied Study Using the Analytic Hierarchy Process to Translate Common Verbal Phrases to Numerical Probabilities. *Journal of Behavioral Decision Making*, (1997), **10(2)**, 133–150.
[27] K. H. Teigen. When Equal Chances = Good Chances: Verbal Probabilities and the Equiprobability Effect. *Organizational Behavior and Human Decision Processes*, (2001), **85(1)**, 77–108.
[28] G. Villejoubert, L. Almond, L. Alison. Interpreting claims in offender profiles: the role of probability phrases, base-rates and perceived dangerousness. *Applied Cognitive Psychology*, (2009), **23(1),** 36–54.
[29] D. Sharpe. Of apples and oranges, file drawers and garbage: why validity issues in meta-analysis will not go away. *Clinical Psychology Review*, (1997), **17(8)**, 881–901.