Sometimes, often, and always: Exploring the vague meanings of frequency expressions

Franziska Bocklisch · Steffen F. Bocklisch · Josef F. Krems

Abstract The article describes a general two-step procedure for the numerical translation of vague linguistic terms (LTs). The suggested procedure consists of empirical and model components, including (1) participants' estimates of numerical values corresponding to verbal terms and (2) modeling of the empirical data using fuzzy membership functions (MFs), respectively. The procedure is outlined in two studies for data from $N = 89$ and $N = 109$ participants, who were asked to estimate numbers corresponding to 11 verbal frequency expressions (e.g., *sometimes*). Positions and shapes of the resulting MFs varied considerably in symmetry, vagueness, and overlap and are indicative of the different meanings of the vague frequency expressions. Words were not distributed equidistantly across the numerical scale. This has important implications for the many questionnaires that use verbal rating scales, which consist of frequency expressions and operate on the premise of equidistance. These results are discussed for an exemplar questionnaire (COPSOQ). Furthermore, the variation of the number of prompted LTs (5 vs. 11) showed no influence on the words' interpretations.

Keywords Translation procedure · Linguistic terms · Frequency expressions · Vagueness · Fuzzy membership functions · Verbal rating scale

F. Bocklisch (✉) · J. F. Krems
Department of Psychology, Chemnitz University of Technology,
09107, Chemnitz, Germany
e-mail: franziska.bocklisch@psychologie.tu-chemnitz.de

S. F. Bocklisch
Department of Automation, Chemnitz University of Technology,
09107, Chemnitz, Germany

Since the 1960s, researchers in different scientific areas have sustained an interest in studying the relationship between verbal and numerical expressions—particularly, probability words and quantifiers (Bocklisch, Bocklisch, & Krems, 2010; Dhami & Wallsten, 2005; Lichtenstein & Newman, 1967; Teigen & Brun, 2003). Moreover, expressions of intensity or frequency of occurrence (e.g., *sometimes* or *often*) are of interest with regard to their wide application in questionnaires. Several studies consistently showed that people prefer to use words instead of numbers to indicate their opinions and uncertainty (e.g., Wallsten, Budescu, Zwick, & Kemp, 1993). Even experts such as doctors or lawyers frequently use qualitative rather than quantitative terms to express their beliefs, on the grounds that words are more natural and are easier to understand and communicate. Words are especially useful in most everyday situations when subjective belief or uncertainty cannot be precisely verbalized in quantitative terms. Therefore, while it may be more natural for people to use language to express their beliefs, it is also potentially more advantageous to use numerical estimates: Their standard interpretation renders them easily comparable, and they form the basis of calculations and computational inferences. Accordingly, many researchers have developed translation procedures (e.g., Beyth-Marom, 1982; Bocklisch et al., 2010; Budescu, Karelitz, & Wallsten, 2003) and have established numerical equivalents for common linguistic expressions (for a broader literature review, see Teigen & Brun, 2003). One outcome of these efforts is that linguistic terms have often been conceptualized as fuzzy sets and mathematically described using fuzzy membership functions (MFs; Budescu et al., 2003; Zadeh, 1965; Zimmer, 1984).

Figure 1 shows an example of the fuzzy MF for the linguistic term *probable* reported by Bocklisch et al. (2010). The function's shape and position represent the vague
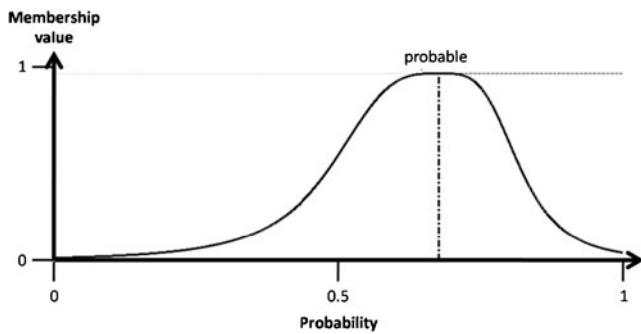
Fig. 1 Example of a fuzzy membership function for the linguistic term *probable* (see Bocklisch, Bocklisch, & Krems, 2010)

meaning of *probable* on a 0–1.0 probability scale. The numerical probabilities occurring between approximately $P = .6$ and $P = .75$ show the highest membership values and, therefore, are most representative and describe the meaning of *probable* best. Because the vague linguistic term has no sharp boundary, the membership values for the other numerical probabilities decrease continuously from the function's peak. Hence, they are less representative of the meaning of *probable*.

The two studies presented herein support the objectives of our article. First, we present a general two-step procedure for the translation of linguistic expressions into numbers and show that this is a methodological innovation. To this end, in study 1, we outline the method exemplarily for verbal frequency expressions. Second, we apply the procedure to the field of verbal rating scales and, thereby, test and construct scales with nearly equidistant response categories. In study 2 we use the verbal response scale of the Copenhagen Psychosocial Questionnaire (COPSOQ; Kristensen, Hannerz, Høgh, & Borg, 2005) as an example. In the Conclusions section, we summarize and outline implications of our results, which include recommendations for the construction of verbal rating scales. Additionally, we discuss interesting future prospects using fuzzy methodology.

Translation procedure as a methodological innovation

The translation procedure is composed of (1) a direct empirical estimation method that yields data from participants who assign numbers to presented words and (2) a fuzzy approach for the analysis of data resulting in parametric MFs of potential type (Bocklisch & Bitterlich, 1994). Our method differs from existing approaches, and the proposed MF type offers advantages over other MF concepts. First, the direct estimation method is very frugal, efficient, and easy to use for yielding empirical data from decision makers. Moreover, our method conserves resources (e.g., as compared with

Budescu et al., 2003) because only three numbers per verbal expression are required for estimation. In our opinion, this is an important criterion regarding potential fields of application (such as medicine) where expert knowledge is crucial but difficult to obtain or expensive. In contrast, Budescu and colleagues proposed a multistimuli method where participants viewed one phrase and 11 probability values (0, .1, . . . , .9, 1) and then judged the degree to which the phrase accurately described each probability. Thus, while these judgments were used to create individualized MFs, they were only partly defined according to the 11 numerical probability values reported by participants. Second, our parametric MFs are defined for a sample or specific population so that a generalized model for the vague linguistic expressions that are suitable for a group of people is obtained. It is a well-known fact that the interindividual variability of estimates is large (Teigen & Brun, 2003). Therefore, if group MFs are fitted, it is necessary to consider variability and potential contradictions in the estimation behavior of participants. The presented MF approach takes this into account by using parameters (see the Method section). Furthermore, we argue that continuous modeling of group MFs of verbal expressions is useful in that it serves as a flexible basis for further calculations. Additionally, such modeling is easily implemented in a variety of existing models or applications, such as decision support systems (Boegl, Adlassnig, Hayashi, Rothenfluh, & Leitich, 2004).

In Bocklisch et al. (2010), the suggested translation method was outlined for verbal probability expressions (e.g., *probable*). The proposed general procedure can be broadly applied to other linguistic terms. In this article, we present the results of two studies. Study 1 included 11 expressions indicative of frequency of occurrence (e.g., *occasionally*) with regard to the potential interest of different research areas and applications that apply verbal rating scales with frequency expressions. After presenting the method, results are discussed with respect to the selection of frequency terms considered appropriate for verbal rating scales in questionnaires. Study 2 employed the translation procedure to explore the COPSOQ response scale in more detail.

Application in verbal response scales

In psychology and the social sciences, many research questions are addressed by directly interrogating participants with the help of questionnaires. Often, responses to presented questions are given by choosing a category of a related verbal answering scale. Although such data collection is determined directly by the verbal categories of the scales, little systematic research has been done (Rohrmann, 1978), as compared with the construction of questionnaire items. Spector (1976)

summarized the consequences of how response categories are commonly selected: "This selection is often made on no more solid basis than habit, imitation, or subjective judgment. Yet the equal interval properties of the response continuum is assumed even though this assumption may, in fact, be false.... When faced with a scale of unequal intervals, subjects sometimes complain of a difficulty in making responses because some adjacent choices are closer together than others. To eliminate this problem, equal interval response categories should be used" (p. 374). Here, we show that our proposed translation procedure can serve as a useful basis for testing and constructing verbal rating scales and determining equidistant verbal response categories.

For the selection of frequency terms, three main criteria are suggested: equidistance, percentage of correct reclassifications, and discriminatory power of the MFs. First, frequency words should be distributed equidistantly along the numerical scale so that data can be interpreted as having interval-scale properties and, therefore, further statistical analyses are feasible. Generally, verbal rating scale categories are assumed to have rank order, but the distance between intervals is not necessarily equal (Jamieson, 2004). That is, verbal rating scale responses comprise ordinal- but not interval-level data, and this precludes the application of parametric statistical analyses. It is common practice to apply mathematical operations, such as multiplication or division (necessary for the calculation of means, etc.) to such data, although these operations are not valid for ordinal data. Moreover, employing inappropriate statistical techniques may lead to the misinterpretation of results and to incorrect conclusions.

Second, the percentage of correct reclassifications—that is, how many original data points were reclassified correctly according to the frequency expression to which they originally belonged—gives information about the discriminability and steadiness of the words' meanings. Third, the criterion of discriminatory power reveals whether MFs differ considerably or not. On the basis of this measure, it is possible to conclude whether the meanings of LTs are interpreted similarly or differently by study participants.

In study 2, fuzzy MFs for the scale of an example questionnaire—namely, the COPSOQ (Kristensen et al., 2005)—are discussed. The COPSOQ is a free screening instrument for evaluating psychosocial factors at work, including stress and employee well-being, as well as selected personality factors. The questionnaire consists of five frequency words: *almost never*, *infrequently*, *sometimes*, *often*, and *always*. We constructed three response scales with alternative frequency expressions and empirically tested an alternative scale consisting of *never*, *sometimes*, *in half of the cases*, *often*, and *always*. We hypothesized that the distance between each of the alternative response labels is nearly equal

and compared results of both scales (original vs. alternative COPSOQ).

Study 1

Method

Two-step translation procedure

Here, we present details of the two-step translation procedure for the numerical translation of verbal frequency expressions. First, the estimation technique and method applied in the empirical study are outlined. Thereafter, fuzzy analysis and MFs are specified.

Step One: empirical investigation

Participants Eighty-nine undergraduate students (9 males) at Chemnitz University of Technology with an average age of 21.5 years ($SD = 2.7$) took part in the study. Four persons stated that they did not understand the task and were therefore excluded from further data analyses.

Materials and procedure The survey instrument was a paper questionnaire and consisted of two parts. In the first part, participants were asked to consider their workload and related requirements that their course of study imposed on them. Then they were asked to answer the following three questions of the COPSOQ (the original material was presented in German): (1) Is it always necessary to work at a rapid pace? (2) Is your work unevenly distributed such that it piles up? (3) How often do you not have enough time to complete all of your work tasks? An explanation as to how the paper questionnaire should be filled out followed, and participants were then asked to assign three numerical values to each of the 11 exemplars of frequency expressions (see translations from the original German in Table 1). Words were chosen according to their frequent usage in questionnaires and in daily communication and on the basis of former research (e.g., Rohrmann, 1978). Three numerical values were estimated: (1) the *typical value* that best represented the given frequency word, (2) the *minimal value*, and (3) *maximal value* that corresponded to the given verbal expression. The semantic meaning of the words can be characterized as follows: The first value identifies the most typical numerical equivalent for the word, whereas other values indicate lower and upper boundaries of the verbal frequency expression. Participants were instructed to give their estimates in frequency format (e.g., Is it *hardly ever* necessary to work at a rapid pace means "in X of 100 work tasks/cases"). We used this format because it is a natural mode of representing information and it turned out

Table 1 Descriptive statistics for the estimates (typical values)

Frequency Expressions (Original German)	Mean	SD	Skew	Kurtosis
Never (nie)	1.37	2.23	1.49	1.26
Almost never (fast nie)	8.31	5.03	0.78	0.72
Infrequently (selten)	18.52	6.36	0.55	0.49
Occasionally (gelegentlich)	28.92	12.23	0.09	−0.71
Sometimes (manchmal)	33.13	10.96	0.44	0.43
In half of the cases (in der Hälfte der Fälle)	50.14	1.21	2.89	21.17
Frequently (häufig)	66.11	15.43	−1.31	2.32
Often (oft)	69.66	12.91	−1.53	3.56
Most of the time (meistens)	75.46	9.05	−0.45	0.01
Almost always (fast immer)	88.11	9.46	−3.48	17.44
Always (immer)	97.46	6.17	−5.83	42.73

that encoding and estimating information in frequency format is easier than in probability or percentage form (Gigerenzer & Hoffrage, 1995; Hoffrage, Lindsey, Hertwig, & Gigerenzer, 2000).

Step two: Fuzzy analysis

Fuzzy membership functions MFs are truth value functions. The membership value (μ) represents the value of the truth that an object belongs to a specific class (e.g., the numerical frequency that 70 of 100 cases belong to the class frequency expression *often*). For the analysis of empirical data provided by the 85 participants, a parametric MF of the potential type (Bocklisch & Bitterlich, 1994; Hempel & Bocklisch, 2009) was used (see Fig. 2).
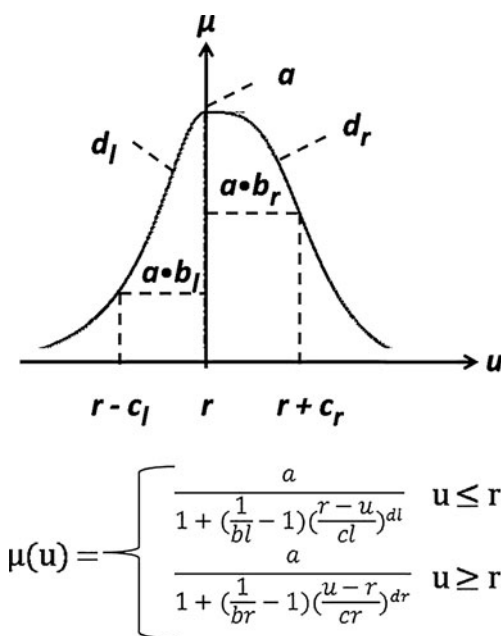

$$\mu(u) = \begin{cases} \dfrac{a}{1 + (\frac{1}{bl} - 1)(\frac{r-u}{cl})^{dl}} & u \leq r \\[4mm] \dfrac{a}{1 + (\frac{1}{br} - 1)(\frac{u-r}{cr})^{dr}} & u \geq r \end{cases}$$

Fig. 2 Parametric membership function of potential type

This function is based on a set of eight parameters: r marks the position of the mean value of the empirical estimates of the *typical value*, while a represents the maximum value of the MF. Regarding class structure, a expresses class *weight* in the given structure (we used $a = 1$ for all classes in this investigation, such that all frequency terms were weighted equally). The parameters c_l and c_r characterize left- and right-sided expansions of the class and, therefore, mark the range of the class, in a crisp sense. In addition to the mean of typical estimates (M_{typ}), the means of minimum (M_{min}) and maximum (M_{max}) correspondence values estimated by participants were used for the calculation: $c_l = M_{typ} - M_{min}$ and $c_r = M_{max} - M_{typ}$. A special feature of this function type is that there is no intersection with the x-axis (μ is always >0). This characteristic is founded on the assumption that sample estimates are not representative of the whole population; therefore, no definite end-points are defined. The parameters b_l and b_r assign left- and right-sided membership values at the boundaries of the function. Therefore, b_l and b_r represent border membership, whereas d_l and d_r specify continuous decline of the MF starting from the class center and are denoted as representative of a class. The d parameters determine the shape of the function and, hence, the fuzziness of the class. The b and d parameters were calculated from the distribution of the empirical data using Fuzzy Toolbox software (Bocklisch, 2008), which is specialized for fuzzy analyses and modeling of MFs.

In contrast to the nonparametric individualized MF approaches of Wallsten, Budescu, Rapoport, Zwick, and Forsyth (1986) and Budescu et al. (2003), we fit group MFs to obtain a generalized model of a sample or certain population of participants. Furthermore, our MFs are defined continuously, such that, in addition to the expansions of the class (c parameters), the MFs' shape (d parameters) carries information about the distribution of the empirical estimates. This is an advantage insofar as potential contradictions between participants' estimates are considered.

In contrast, a triangular MF type describes the graded interval between $\mu = 0$ and $\mu = 1$ with a rather arbitrary linear model and, thus, does not account for the empirical data provided by many individuals. On the level of individual estimates, a triangular MF would model the data appropriately, but on the level of a certain sample or population, this is not the case. Additional parameters are needed to model the expansion (c) and the distribution of the estimates (d), as well as the membership value at the border of the function (b), which is by definition always >0. A continuous variation of MFs, ranging from highly fuzzy to crisp, is available through this parametric function type. It also allows for asymmetry in fuzzy classes by providing individual parameters for the left- and right-hand branches of the function. As the results of former research show (Bocklisch et al., 2010; Budescu et al., 2003), many verbal expressions are best described by asymmetric MFs. Therefore, we expect this feature to be especially important for the present study.

Results

We first present the descriptive statistics of the data set. Thereafter, the fuzzy MF procedure is specified. In our opinion, it is valuable to present both results for purposes of completeness and comparison, even though we favor the latter approach. It is important that the two approaches be understood independently. Moreover, fuzzy analysis and modeling of the MFs, by definition, do not refer to the background of probability theory and statistics. Although some parameters of our MF type can be interpreted statistically in this case (e.g., r values are equal to the arithmetic mean), an MF is not a probability density function, and conventional requirements (i.e., the integral of the variable's density is equal to 1) are not valid. A more general comparative discussion of the statistical and fuzzy approaches is provided in Singpurwalla and Booker (2004).

Descriptive statistics

Table 1 shows the typical values that corresponded to the frequency expressions presented. Minimum and maximum estimates of the semantic meaning of linguistic terms were necessary for modeling the MFs (c parameters). Hence, they are not presented here.

At first glance, the results show that frequency expressions are distributed almost over the entire numerical frequency scale with varying distances, ranging from *never* ($M = 1.37$) to *always* ($M = 97.46$). Clearly, the 11 expressions are divided into three frequency categories: lower and higher frequency categories, which refer to the middle point of the scale ($M = 50$), and a medium frequency category consisting of one LT (*in half of the cases*: $M = 50.14$). The first 5

expressions (ranging from *never* to *sometimes*) are characterized by means less than $M = 35$ and, therefore, belong to the lower frequency group, whereas the remaining expressions (ranging from *frequently* to *always*) show mean values larger than $M = 65$ and belong to the higher frequency category. Between the expressions *sometimes* and *in half of the cases* and between *in half of the cases* and *frequently*, there are intervals measuring approximately 15. These are the largest two intervals among all the intervals between the LTs. Similar findings were reported by Bocklisch et al. (2010) for verbal probability expressions, which are also split according to three categories (low, medium, and high probability). Standard deviation (*SD*) values show a systematic pattern: Frequency expressions near the borders of the numerical frequency scale have smaller *SD*s. Starting with the minimum of the verbal scale (*never*: $SD = 2.23$), the SDs increase up to midscale, reaching their highest values with the words *occasionally* ($SD = 12.23$) and *sometimes* ($SD = 10.96$), as well as *frequently* ($SD = 15.43$) and *often* ($SD = 12.91$), and subsequently decrease again (*always*: $SD = 6.17$). Again, the frequency expression that covers the middle of the scale (*in half of the cases*: $SD = 1.21$) is an exception, because its SD is the smallest one. By tendency, skews are higher at the borders of the verbal scale. Expressions belonging to the lower category (e.g., *never*) are slightly skewed to the right, and in the higher category (e.g., *always*), they tend to be skewed to the left. Kurtosis values are considerably higher for the expressions *in half of the cases*, *almost always*, and *always*, while values for the other frequency expressions are almost normally distributed (i.e., kurtosis = 0 according to the SPSS software's definition). These findings are consistent with results reported by Bocklisch et al. (2010) as well as Budescu et al. (2003) that investigated verbal probability expressions.

Fuzzy analysis

Figure 3 shows the MFs for the 11 verbal frequency expressions. The representative values (r) indicating the highest memberships are identical to the reported means in Table 1. Obviously, the functions differ in shape, symmetry, overlap, and vagueness. The functions for the verbal frequency expressions at the borders of the scale, *never* and *always*, are narrower than those in the middle, such as *sometimes* or *often*, which is in accordance with reported *SD*s and kurtosis values. Most of the functions are slightly asymmetric and are clearly not distributed equidistantly along the scale. Some (neighbor) functions overlap to a large extent (e.g., *occasionally* and *sometimes*), while others are quite distinct (e.g., *in half of the cases* and *frequently*).

The area of MF overlap A_{ov} (see Fig. 4, gray area) is informative about the similarity of the words' meanings. Overlap is defined as the surface imbedded by the MFs and
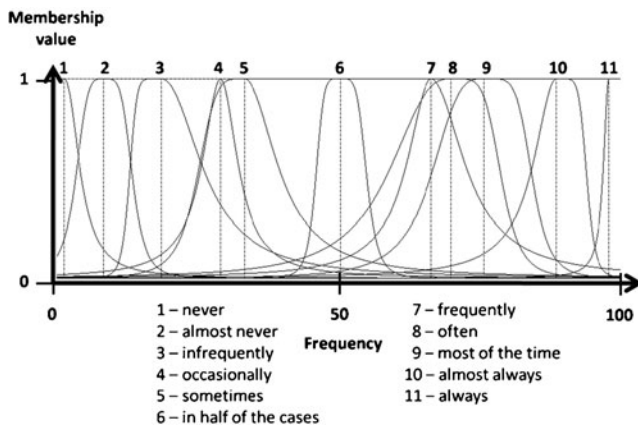
Membership
value

Fig. 3 Membership functions of the 11 verbal frequency expressions

the x-axis. One important characteristic of our parametric potential MF type is that the function has no points of intersection with the x-axis and, therefore, the surface integral is infinite. Additionally, the function type has no general integral solution. Hence, the surface covered by the function (in a certain range) can only be approximated, which is done with the help of Fuzzy Toolbox software (Bocklisch, 2008) and operates as follows. The range of the MFs is identified: Here, the minimum is 0 and the maximum is 100 according to the numerical frequency scale. Thereafter, μ_{min} is calculated numerous times using a high sampling rate with equidistant sample points along the numerical scale. Then the area of overlap A_{ov} is determined by adding up the products of the sampling distance and μ_{min} values for the whole number of sampling points. Thereafter, areas covered by MF1 and MF2 (A_{MF1} and A_{MF2}) are defined using the same procedure. A standardized quotient (ov) of the overlapping area of the MFs (A_{ov}) is obtained by calculating the arithmetic mean: $ov = 0.5 \times [(A_{ov} : A_{MF1}) + (A_{ov} : A_{MF2})]$.

The ov is used to define the discriminatory power (dp) between two MFs: $dp = 1 - ov$ (Bocklisch, 2008). The dp is standardized taking values from 0 (MFs are identical)
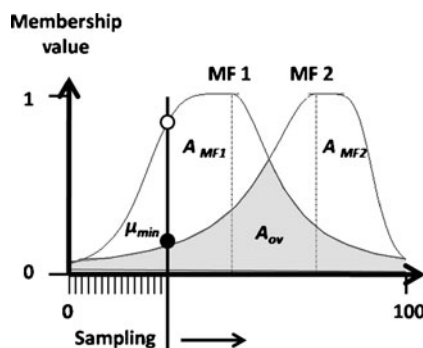

Fig. 4 Approximation of the discriminatory power of two membership functions

to 1 (no overlap at all). Hence, the larger the overlap (e.g., occasionally and sometimes), the smaller the dp and the more similar the meanings of the verbal expressions are. The ov of the MFs in Fig. 4 is approximately .37 which corresponds to dp = .63. Table 2 shows dp values for the 11 LTs.

If dp values are greater than or equal to .7, then MFs (and LTs) are interpreted as being considerably different, because the area of shared overlap is less than 30%. This is the case for a lot of LTs (see Table 2), except for infrequently and occasionally (dp = .46), occasionally and sometimes (dp = .25), frequently and often (dp = .19), often and most of the time (dp = .32), frequently and most of the time (dp = .38), and most of the time and almost always (dp = .69). Most of these LT pairs are direct "neighbors."

The COPSOQ answer scale (Kristensen et al., 2005) consists of five frequency expressions: almost never, infrequently, sometimes, often, and always. Figure 5 shows the MFs of the verbal rating scale utilized in the COPSOQ (upper left corner) and three proposed alternative scales that are almost equidistant, consisting of four and five frequency expressions.

In the original COPSOQ scale, the distances between the representative values vary. The LTs almost never and infrequently have approximately the same distance (10.21) as infrequently and sometimes (14.61), but the words sometimes and often (36.53), as well as often and always (27.8), are separated by a greater distance. Therefore, this scale is not equidistant. Furthermore, no verbal term is associated with the middle of the scale, which indicates a frequency of occurrence of approximately 50 out of 100. That is, such a term is unavailable, even to participants who should wish to express this frequency.

The interpretation of verbal frequency scales as interval scales relies on the premise of equidistance (Jamieson, 2004). While authors of the COPSOQ may have wanted the frequency words to be distributed as shown in Fig. 5, such a distribution is rather unlikely, for two reasons: First, if a middle category is not intended, an even number of LTs is usually chosen for a verbal response scale. Second, a scale that combines highly similar words (such as almost never and infrequently) with highly discriminatory terms (e.g., often and always) seems to be inconsistent.

To remedy this problem, we propose three scales that meet the criterion of equidistance quite well (see Fig. 5): first, two 5-point scales consisting of the frequency terms never, sometimes, in half of the cases, often, and always (alternative COPSOQ I) and almost never, sometimes, in half of the cases, often, and almost always (alternative COPSOQ II) and, second, a 4-point scale with the expressions almost never, sometimes, often, and almost always (alternative COPSOQ III). The frequency expressions for these scales were chosen according to results presented in

Table 2 Discriminatory power values for the 11 MFs

LT		1	2	3	4	5	6	7	8	9	10	11
1	Never	0										
2	Almost never	.83	0									
3	Infrequently	.98	.86	0								
4	Occasionally	.94	.91	.46	0							
5	Sometimes	.98	.97	.59	.25	0						
6	In half of the cases	1.00	1.00	.93	.94	.83	0					
7	Frequently	.98	.98	.88	.90	.84	.86	0				
8	Often	.98	.98	.87	.88	.82	.77	.19	0			
9	Most of the time	.99	1.00	.93	.95	.91	.93	.38	.32	0		
10	Almost always	.98	.99	.93	.94	.91	.97	.74	.82	.69	0	
11	Always	1.00	1.00	.98	.99	.98	1.00	.94	1.00	.99	.97	0

Table 2 and Fig. 3. Both 5-point scales (alternative COPSOQs I and II) are distributed almost equidistantly, do not overlap to a great extent (see *dp* values in Table 2), and are almost symmetric in shape. However, they differ according to their psychological width, which ". . . refers to the extent of the psychological continuum suggested by the rating labels" (Lam & Stevens, 1994, p.142). Therefore, alternative COPSOQ I is wider, because the LTs at the borders of the scale approximate the numerical endpoints (*never*, M = 1.37; *always*, M = 97.46) and, hence, mark a wider psychological continuum than the LTs of alternative COPSOQ II (*almost*

never, M = 8.31; *almost always*, M = 88.11). The 4-point alternative COPSOQ III (see Fig. 5, lower left) is also nearly equidistant, where MFs are highly distinct and the middle of the scale is not covered.

In addition to the criteria of equidistance, symmetry, and overlap of the MFs' distribution, the percentage of correct reclassifications of the participants' original estimates is informative of the quality of the scales. For the reclassification task, the original data were used and reassigned to the MFs. Basically this is done by using a participant's typical estimate for a certain verbal expression and entering it into the
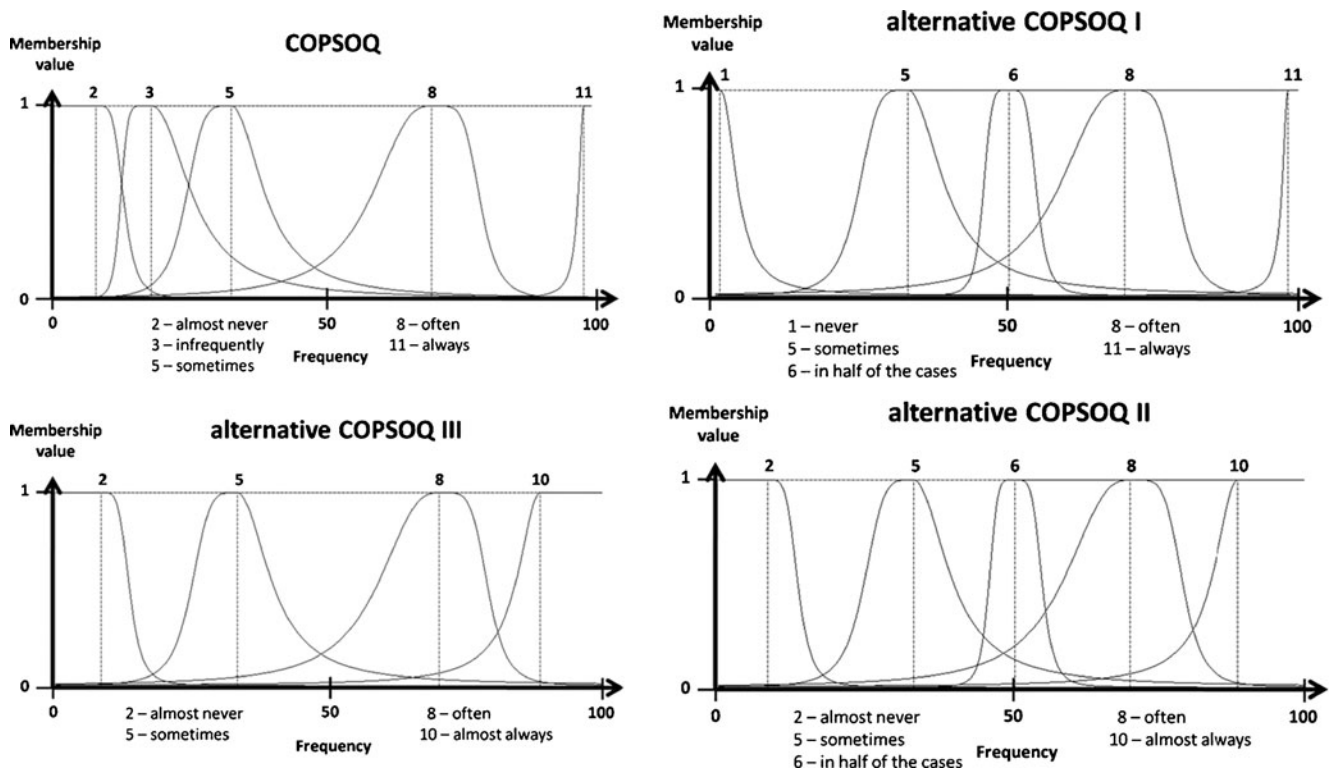

Fig. 5 Membership functions of the original COPSOQ and alternative COPSOQ (I-III) response scales

equations of all MFs (see Fig. 2) as u. Then the membership values (μ) can be calculated. Therefore, 11 membership values (i.e., for the 11 MFs of the 11 frequency expressions) are generated for one data point (i.e., estimate of a respondent). Among these, the highest membership value indicates the frequency word to which the estimate is reclassified. The reclassification is correct if this frequency word is the same as the one for which the estimate was originally given. The reclassification step was done with the help of Fuzzy Toolbox software (Bocklisch, 2008). Table 3 shows reclassification results obtained by counting the number of original data points correctly reclassified according to the frequency expression to which they originally belong.

For the original scale consisting of 11 frequency expressions, the correct reclassification percentages lie between 1.18% for *occasionally* (only 1.18% of the typical estimates for *occasionally* were reclassified as belonging to *occasionally*, and the other 98.82% were erroneously reclassified as belonging to other frequency expressions) and 98.82% for *in half of the cases* (nearly all estimates for *in half of the cases* were reclassified as belonging to *in half of the cases*). The mean percentage of reclassification for this scale (M = 56.35) is rather low, which is mainly due to the large overlap of the MFs of the frequency expressions (see Fig. 3). The original COPSOQ scale (M = 79.99) and all alternative scales (M > 85.3) with four to five linguistic terms have higher mean percentages of correct reclassification. Hence, the more terms that are included in a scale, the lower the reclassification percentages will be, due to the similarity of the words' meanings that can be observed in the overlap of the MFs. In summary, all suggested alternative COPSOQ scales showed better reclassification results and were nearly equidistant, as compared with the original COPSOQ scale. To optimize all criteria, it would be advisable to choose the alternative COPSOQ I with the five frequency expressions *never*, *sometimes*, *in half of the cases*, *often*, and *always*.

Discussion

In study 1, we outlined a general procedure for the translation of verbal expressions based on empirical estimates and using fuzzy MFs for modeling. The results (see Table 1 and Fig. 3) showed that the MFs of frequency expressions at borders of the numerical scale (i.e., *never* and *always*) showed less vagueness than did midscale expressions (i.e., *often* and *sometimes*), suggesting that they more clearly reflected the given expression. This was also found for probability expressions (Bocklisch et al., 2010) that differed even more in vagueness when midscale terms and boundary terms are compared. The LT *in half of the cases* is an exception (SD = 1.21; see MF in Fig. 3): Its meaning is rather crisp with regard to other frequency expressions in the middle of the scale and as compared with the midscale probability LTs *thinkable* (SD = 20.24) and *possible* (SD = 21.60) in Bocklisch et al. (2010). This could be due to the relatively "precise" meaning of the word "half."

The *dp* values (see Table 2) and percentages of correct reclassification (see Table 3) were introduced as means for measuring the disparity and steadiness of the MFs. Hence, a differentiated evaluation of the MFs is possible, and conclusions concerning the meaning of the modeled LTs are straightforward. For a few MFs, *dp* values are rather low, and therefore, the meanings of the corresponding LTs are very similar. However, most of the words are distinct. The percentages of correct reclassification are very high for *never* (81.18), *in half of the cases* (98.82), and *always* (91.57), which supports the idea that these LTs are more precise in their meanings.

The emerging categories, low, middle, and high frequencies, may be due to the actual sample of verbal expressions. It would be interesting to determine whether the estimation of more or fewer LTs would lead to the

Table 3 Percentages of correct reclassification

Frequency Expressions (Original German)	11 LTs	COPSOQ	Alternative COPSOQ I	Alternative COPSOQ II	Alternative COPSOQ III
Never (nie)	81.18	–	100	–	–
Almost never (fast nie)	68.24	83.53	–	92.94	92.94
Infrequently (selten)	67.06	67.06	–	—	–
Occasionally (gelegentlich)	1.18	–	–	–	–
Sometimes (manchmal)	56.47	60.00	76.47	76.47	77.65
In half of the cases (in der Hälfte der Fälle)	98.82	–	98.82	98.82	–
Frequently (häufig)	5.88	–	–	–	–
Often (oft)	42.35	91.76	87.06	82.36	87.06
Most of the time (meistens)	49.41	–	–	–	–
Almost always (fast immer)	57.65	–	–	83.53	83.53
Always (immer)	91.57	97.62	97.62	–	–
Mean	56.35	79.99	91.99	86.82	85.3

same categories as those found in this study and in Bocklisch et al. (2010) or not.

Many questionnaires utilize verbal rating scales consisting of verbal frequency expressions. Thus, we exemplarily tested a well-established questionnaire, the COPSOQ, concerning equidistant distribution of its linguistic expressions and the quality of the scale (i.e., percentages of correct reclassification of the original data). It was found that the scale is in need of improvement because it fails to satisfy the criterion of an equidistant distribution. At present, strictly speaking, the scale cannot be interpreted as having interval level, and hence, further statistical analyses (e.g., the calculation of means for groups of participants) are not appropriate. To solve this problem we proposed three alternative COPSOQ scales with four or five frequency expressions distributed nearly equidistantly (see Fig. 5). The suggested 4-point scale (alternative COPSOQ III) should be employed for research questions where no middle category is intended. Alternatives I and II differ concerning LTs at the borders, and alternative I offers a wider psychological continuum for frequency estimation. Both scales produced positive results for mean reclassification percentages, *dp*s of the MFs, and equidistance and can thus both be applied according to intended utilization. Wyatt and Meyers (1987) found that scales with less extreme endpoints (e.g., alternative COPSOQ II: *almost never* and *almost always*) lead to greater variability in respondents' estimates than do scales with more extreme endpoints (e.g., alternative COPSOQ I: *never* and *always*). However, it is not yet clear whether this finding can be generalized to other words and contexts (Lam & Stevens, 1994).

In summary, we showed that our translation procedure is a methodological innovation and, therefore, has potential for application in research. In study 2 we use the method again, exploring the COPSOQ scale in greater detail. That is, one could argue that the total number of frequency expressions influences the resulting MFs. If this were the case, it might be inappropriate to draw conclusions from a study that presented 11 LTs to a scale (COPSOQ) that consisted of only 5 LTs. Therefore, in study 2, we presented the 5 LTs and compared the results with those of study 1. Additionally, we manipulated scales of the original COPSOQ and alternative COPSOQ I, which allowed us to test whether our conclusions based on the MFs in study 1 were indeed correct.

Study 2

Method

Participants One hundred nine undergraduate students (19 males) of Chemnitz University of Technology with an average age of 23.4 years (SD = 3.3) took part in the study. Fifteen persons did not understand the task and were therefore excluded from further analyses.

Materials and procedure The paper questionnaire employed in study 2 was identical to that used in study 1, except that the number of presented frequency expressions differed (study 1, 11 LTs vs. study 2, 5 LTs). Again, participants first answered three questions of the COPSOQ. One group of participants ($N = 51$) received the original COPSOQ response scale (*almost never*, *infrequently*, *sometimes*, *often*, and *always*), while the other group ($N = 42$) obtained an alternative COPSOQ answering scale (*never*, *sometimes*, *in half of the cases*, *often*, and *always*). In the second part, the study 1 translation procedure was also used to translate the five frequency expressions.

Results

Descriptive statistics

Table 4 shows the descriptive results of the typical values that corresponded to the frequency expressions of the original and alternative COPSOQ scales (middle and right columns), as well as the results of study 1 (left column; see also Table 1) for purposes of comparison.

For the LTs *sometimes*, *often*, and *always*, a direct comparison between all conditions is possible. In sum, mean values for *often* and *always* are very similar. The largest difference is 5.3 between *always* in the context of 11 LTs and *always* in the original COPSOQ scale using 5 LTs. For *sometimes*, the original COPSOQ ($M = 41.08$) stands out, as compared with the other conditions (alternative COPSOQ, $M = 29.0$ and the 11-LT version, $M = 33.13$). The differences between conditions for *never* and *in half of the cases* (11 LTs vs. 5LTs. alternative COPSOQ) as well as for *almost always* and *infrequently* (11 LTs vs. 5LTs, original COPSOQ) are also rather small. The SDs are comparable in size between groups for a certain LT, except *always* (original COPSOQ: $SD = 19.04$), which has a larger SD than the other conditions.

Fuzzy analysis

Figure 6 shows the resulting MFs for the five verbal frequency expressions of the original versus alternative COPSOQ response scales in the context of 5 LTs vs. 11 LTs (see also Fig. 5).

In the alternative scale version (5 LTs), the verbal terms at the borders (*never* and *always*) are closer to the borders of the underlying numerical scale, as compared with the original scale (5 LTs). The scales also differ in the extent of

Table 4 Comparison of descriptive statistics of study 1 (11 LTs) versus study 2 (5 LTs) for work context (original vs. alternative COPSOQ answering scales)

Frequency Expressions (Original German)	Study 1		Study 2			
	Mean 11 LTs	SD	Mean 5 LTs (original)	SD	Mean 5 LTs (alternative)	SD
	N = 85		N = 51		N = 42	
Never (nie)	1.37	2.23	–	–	2.71	3.93
Almost never (fast nie)	8.31	5.03	9.80	4.98	–	–
Infrequently (selten)	18.52	6.36	20.31	8.01	–	–
Sometimes (manchmal)	33.13	10.96	41.08	12.38	29.00	11.94
In half of the cases (in der Hälfte der Fälle)	50.14	1.21	–	–	50.24	1.54
Often (oft)	69.66	12.91	73.40	13.10	69.27	17.41
Always (immer)	97.46	6.17	92.16	19.04	96.74	4.97

the MFs' overlaps. For instance, in the original COPSOQ, the overlaps occurring at border terms are larger, and in the alternative version, midscale terms overlap more. The distribution of MFs is closer to equidistance for the suggested alternative response scale. The functions' shapes of the word *often* are very similar, while the others differ slightly—for instance, in expansion (e.g., the MF for *sometimes* is broader in the alternative scale version). The frequency expression *in half of the cases* marks the middle of the scale. The function's shape is salient; it is

asymmetric, and the left-hand branch is very crisp, as compared with the right-hand branch.

A comparison of frequency expressions between the 5- and 11-LT versions of the original COPSOQ (see Fig. 6, left side) and of the alternative COPSOQ (see Fig. 6, right side) shows a highly similar appearance of MFs in terms of *r*-value positions (equal to the means in Table 4), shapes, and overlaps. MFs tend to be slightly narrower in the 11-LT versions of the two scales, and the border term *always* tends to be more extreme, as compared
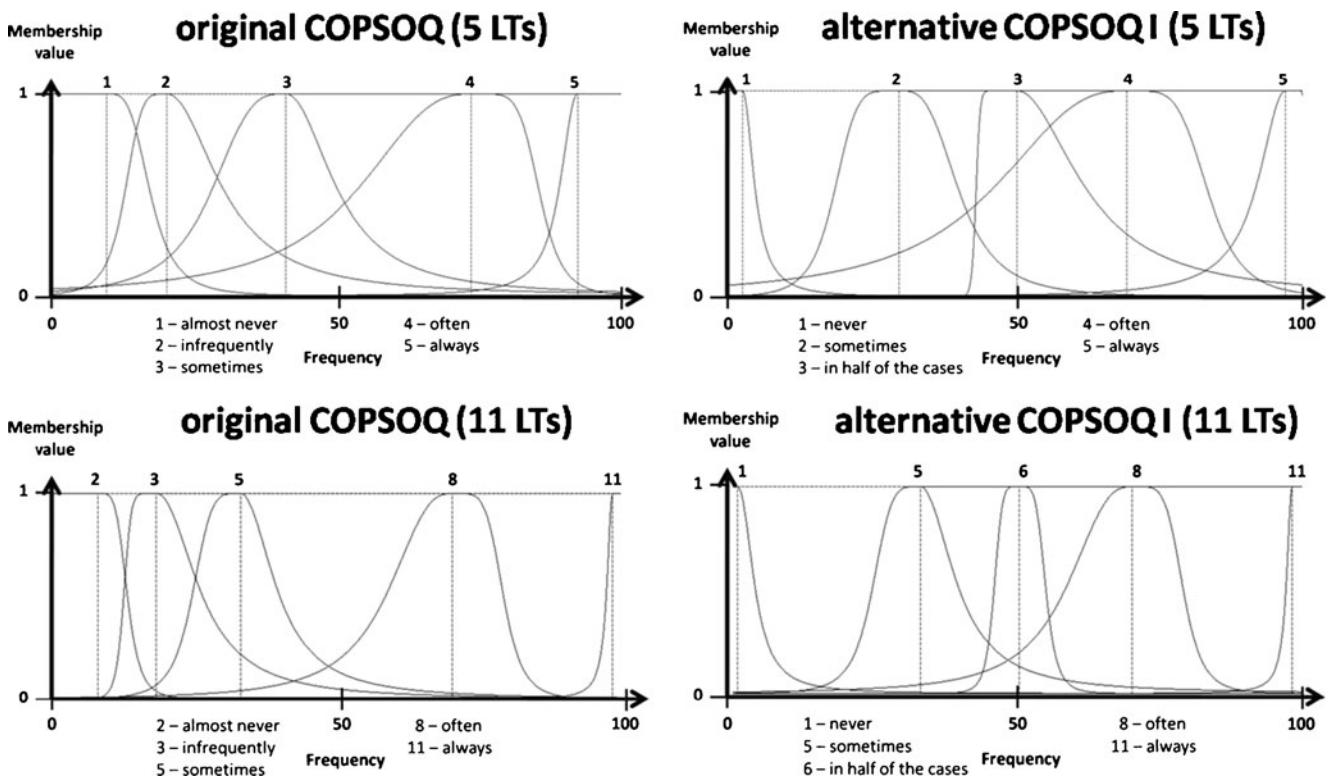

Fig. 6 Membership functions of the verbal frequency expressions of the original versus alternative COPSOQ I response scales for 5 versus 11 LTs

with the 5-LT versions. The frequency expression *in half of the cases* has equal *r* values (5 LTs, $r = 50.24$; 11 LTs, $r = 50.14$), but the MF's shape deviates. In the 5-LT version of the alternative COPSOQ, it is rather fuzzy and asymmetric, whereas in the 11-LT version, it is very crisp and symmetric. For the evaluation of the differences between the 5- and 11-LT versions, again, *dp* values are calculated. Table 5 shows the *dp* values.

For instance, for *sometimes*, the difference between the 5- and 11-LT versions of the original COPSOQ scale is slightly larger ($dp = .29$) than for the 5- and 11-LT versions of the alternative COPSOQ I scale ($dp = .14$). Generally, *dp* values for *never, almost never, infrequently, sometimes, in half of the cases*, and *often* are all rather small ($dps < .49$), which means that the MFs are very similar and overlap in 50% to 90%. However, for *always*, there is a considerable difference between MFs in the alternative COPSOQ I (5 vs. 11 LTs: $dp = .74$), but not for the original COPSOQ (5 vs. 11 LTs: $dp = .53$).

Discussion

Study 2 aimed to clarify (1) whether the suggested alternative response labels (see Fig. 5: alternative COPSOQ I) also have equal distances in the context of 5 LTs and (2) whether the total number of prompted LTs (5 vs. 11) influences the interpretation of frequency words. First, we found that alternative COPSOQ I has nearly equal distances between the response categories (see Table 4 and Fig. 6). Hence, our presented method is generally suitable for application in choosing LTs for answering scales. Second, the resulting *dp* values (see Table 5) show that the total number of prompted LTs seems to have no systematic influence on the words' interpretation, since nearly all MFs are identical to a great extent ($dps < .53$). There is only one considerable difference: MFs of *always* (alternative COPSOQ I) are distinct ($dp = .74$). That is, *always* in the 5-LT version is broader and covers more of the numerical frequency scale than *always* in the 11-LT version does.

Table 5 Discriminatory power values for original and alternative COPSOQ I scales (5 vs. 11 LTs)

LT	Original	Alternative
Never	–	.49
Almost never	.42	–
Infrequently	.14	–
Sometimes	.29	.14
In half of the cases	–	.30
Often	.18	.23
Always	.53	.74

Nevertheless, the difference is rather small, because the criterion value of $dp > .7$ is just met. Accordingly, this tendency is also the case for *always* in the original version (see Fig. 6, left side). Our results show that the number of prompted LTs has no considerable influence on the interpretation of the LTs meanings, although there are, at least to some extent, small differences between the MFs depending on the number of LTs presented (see also Table 4).

Conclusions

This article presents a general two-step procedure for the numerical translation of linguistic terms that are exemplars of frequency expressions. In two studies, we showed that the presented procedure is a methodological innovation and can serve as basis for choosing LTs for applications such as questionnaires. In study 1, the procedure was presented for 11 frequency expressions. First, three numerical values for each linguistic term (i.e., most typical, minimal, and maximal correspondence values) were estimated. Second, the resulting data were modeled using the parametric MFs of the potential type. While most alternative procedures are more costly (Budescu et al., 2003) or are not based on empirical estimates (Boegl et al., 2004), our approach is very frugal and efficient in terms of data collection.

Results show that the functions are capable of modeling the data in a very efficient way, yielding averaged MFs that describe the LTs continuously along a numerical frequency scale. They also take into account the asymmetry of the empirical data, resulting due to the parameters that model the left- and right-hand branches of the function (e.g., c_l and c_r). MFs with fewer parameters would model the data without considering asymmetry and would, therefore, be less accurate and suitable for the reported data. The *b* and *d* parameters reflect features of the distribution of the empirical estimates and carry information about between-subjects differences. Another advantage of the proposed function type is that the semantic content of parameters can be interpreted at a meta-level. Hence, they render the vague meaning of linguistic terms more tangible. In addition to existing methods (e.g., Boegl et al., 2004; Budescu et al., 2003; Wallsten et al., 1986), this parametric MF approach is an interesting alternative that yields group MFs and contributes to the investigation of vague linguistic terms. Future research would benefit from a comparison of different translation procedures and MF concepts (e.g., individualized MFs vs. group MFs).

In study 2, we explored the COPSOQ scale in detail. Questionnaires are widely used in the social sciences and

humanities to address empirical research questions. We exemplarily tested the COPSOQ questionnaire (see the Results sections of studies 1 and 2) and found that the scale employed in this tool is in need of improvement because its verbal labels fail to satisfy the criterion of an equidistant distribution. At present, this questionnaire scale is ordinal rather than interval level, and therefore, statistical analyses such as the calculation of arithmetic means for groups of participants are not valid. A counterargument might suggest that missing equidistance is compensated for by the conventional visual arrangement of scales. This might, indeed, have an influence on the interpretation of the words' meanings. To clarify this issue, our translation approach may be useful for further studies. We suggest three nearly equidistant verbal frequency scales (see Fig. 5) with four or five frequency expressions as a starting point for such studies.

In constructing verbal response scales, we recommend adapting the context of the cover story according to the topic (e.g., psychology, medicine, or economy), because context is known to influence a word's interpretation (Pepper & Prytulak, 1974; Teigen & Brun, 2003). Additionally, the purpose for which the LTs will be used afterward (e.g., questionnaire or decision support system) should also be considered. Future studies may benefit from choosing estimators from the target population—for example, medical experts or participants in experimental studies. According to the desired psychological width of the response scale, "choosing a scale for a particular application must take into account what needs to be measured" (Wyatt & Meyers, 1987, p. 34).

Different samples of participants and different languages of investigation should also be considered in future studies. We report data from a student sample using German LTs. Although this might limit the generalizability of our results, the presented methodology (translation procedure) is not restricted to a certain sample or language. Therefore, it would be interesting to study how different samples of people (such as experts vs. novices in medicine) interpret LTs and whether or not the meanings of verbal expressions are understood similarly in different languages.

The reported MFs, especially in study 1, show large overlaps, indicating that contiguous expressions are very similar or almost identical in their meanings. It is noteworthy that despite the vagueness of natural language, MFs are a convenient tool for identifying words that are more distinct (i.e., with small overlap) in their meaning than others. The identification of unambiguous and distinct words that can be used for communication is of tremendous importance in areas such as medicine or the military, where misunderstandings could lead to severe consequences. Currently, we are exploring the availability of such distinct words for communication purposes with the help of our MFs. Karelitz and Budescu (2004) devised promising criteria for the conversion of phrases from a communicator's to a recipient's lexicon—for instance, the peak rank order between MFs. Our MF approach could contribute additional criteria to such an approach, such as the mathematical quantification of MF overlaps.

Appendix

Table 6 MF parameters of the 11 frequency expressions (see Fig. 3)

Frequency Expressions (Original German)	r	b_l	b_r	c_l	c_r	d_l	d_r
Never (nie)	1.37	1.00	0.69	1.37	2.20	2.00	2.20
Almost never (fast nie)	8.31	0.63	0.63	4.40	4.60	4.11	4.67
Infrequently (selten)	18.52	0.44	0.66	5.80	6.00	8.32	2.14
Occasionally (gelegentlich)	28.92	0.20	0.21	7.80	7.00	2.00	2.00
Sometimes (manchmal)	33.13	0.69	0.50	7.10	6.80	4.65	2.03
In half of the cases (in der Hälfte der Fälle)	50.14	0.34	0.44	4.90	4.90	5.75	4.87
Frequently (häufig)	66.11	0.18	0.37	10.9	9.00	2.00	2.00
Often (oft)	69.66	0.69	0.63	8.70	8.10	2.70	5.38
Most of the time (meistens)	75.46	0.69	0.63	7.50	7.80	3.33	5.68
Almost always (fast immer)	88.11	0.34	0.50	6.60	5.40	2.00	7.30
Always (immer)	97.46	0.06	1.00	4.20	2.54	2.00	2.30

Table 7 Frequencies of correct reclassification for the 11 LTs (see Fig. 3)

Frequency Expressions (Original German)		1	2	3	4	5	6	7	8	9	10	11
1	Never	69	16	0	0	0	0	0	0	0	0	0
2	Almost never	13	58	14	0	0	0	0	0	0	0	0
3	Infrequently	0	19	57	1	8	0	0	0	0	0	0
4	Occasionally	0	8	33	1	32	11	0	0	0	0	0
5	Sometimes	0	1	20	2	48	13	0	1	0	0	0
6	In half of the cases	0	0	0	0	0	84	0	1	0	0	0
7	Frequently	0	1	1	0	5	8	5	34	27	4	0
8	Often	0	0	2	0	1	5	4	36	31	6	0
9	Most of the time	0	0	0	0	0	2	2	25	42	14	0
10	Almost always	0	0	0	0	1	0	0	3	10	49	22
11	Always	0	0	0	0	0	0	0	0	0	7	76

Table 8 Frequencies of correct reclassification for original COPSOQ (see Fig. 4)

Frequency Expressions (Original German)		2	3	5	8	11
2	Almost never	71	14	0	0	0
3	Infrequently	19	57	9	0	0
5	Sometimes	1	20	51	13	0
8	Often	0	2	2	78	3
11	Always	0	0	0	2	82

Table 9 Frequencies of correct reclassification for alternative COPSOQ I (see Fig. 4)

Frequency Expressions (Original German)		1	5	6	8	11
1	Never	85	0	0	0	0
5	Sometimes	1	0	1	1	82
6	In half of the cases	1	5	6	8	11
8	Often	6	65	13	1	0
11	Always	0	0	1	1	82

Table 10 Frequencies of correct reclassification for alternative COPSOQ II (see Fig. 4)

Frequency Expressions (Original German)		2	5	6	8	10
2	Almost never	79	6	0	0	0
5	Sometimes	6	65	13	1	0
6	In half of the cases	0	0	84	1	0
8	Often	0	3	5	70	7
10	Almost always	0	1	0	13	71

Table 11 Frequencies of correct reclassification for the alternative COPSOQ III (see Fig. 4)

Frequency Expressions (Original German)		2	5	8	10
2	Almost never	79	6	0	0
5	Sometimes	6	66	13	0
8	Often	0	4	74	7
10	Almost always	0	1	13	71

References

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

Bocklisch, S. F. (2008). *Handbook Fuzzy Toolbox. GWT-TUD.* Chemnitz, Germany: Chemnitz University of Technoloogy, Department of Electrical Engineering.

Bocklisch, S. F., & Bitterlich, N. (1994). Fuzzy pattern classification—methodology and application. In R. Kruse, J. Gebhardt, & R. Palm

(Eds.), *Fuzzy systems in computer science* (pp. 295–301). Wiesbaden, Germany: Vieweg.

Bocklisch, F., Bocklisch, S.F., & Krems, J.F. (2010). How to translate words into numbers? A fuzzy approach for the numerical translation of verbal probabilities. In E. Hüllermeier, R. Kruse, & F. Hoffmann (Eds.), *IPMU 2010, Lecture Notes in Artificial Intelligence 6178* (pp. 614–623). Springer.

Boegl, K., Adlassnig, K.-P., Hayashi, Y., Rothenfluh, T. E., & Leitich, H. (2004). Knowledge acquisition in the fuzzy knowledge representation framework of a medical consultation system. *Artificial Intelligence in Medicine, 30,* 1–26.

Budescu, D. V., Karelitz, T. M., & Wallsten, T. S. (2003). Predicting the directionality of probability words from their membership functions. *Journal of Behavioral Decision Making, 16,* 159–180.

Dhami, M. K., & Wallsten, T. S. (2005). Interpersonal comparison of subjective probabilities: Towards translating linguistic probabilities. *Memory & Cognition, 33,* 1057–1068.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102,* 684–704.

Hempel, A.-J., & Bocklisch, S. F. (2009). Parametric fuzzy modelling for complex data-inherent structures. In *Proceedings of the Joint 2009 International Fuzzy Systems Association World Congress and 2009 European Society of Fuzzy Logic and Technology Conference (IFSA-EUSFLAT 2009)* (pp. 885–890). Lisbon.

Hoffrage, U., Lindsey, S., Hertwig, R., & Gigerenzer, G. (2000). Communicating statistical information. *Science, 290,* 2261–2262.

Jamieson, S. (2004). Likert scales: How to (ab)use them. *Medical Education, 38,* 1217–1218.

Karelitz, T. M., & Budescu, D. V. (2004). You Say "probable" and I say "likely": Improving interpersonal communication with verbal probability phrases. *Journal of Experimental Psychology. Applied, 10,* 25–41.

Kristensen, T. S., Hannerz, H., Høgh, A., & Borg, V. (2005). The Copenhagen Psychosocial Questionnaire (COPSOQ)—a tool for the assessment and improvement of the psychosocial work environment. *Scandinavian Journal of Work and Environmental Health, 31,* 438–449.

Lam, T. C. M., & Stevens, J. J. (1994). Effects of content polarization, item wording, and rating scale width on rating responses. *Applied Measurement in Education, 7,* 141–158.

Lichtenstein, S., & Newman, J. R. (1967). Empirical scaling of common verbal phrases associated with numerical probabilities. *Psychonomic Science, 9,* 563–564.

Pepper, S., & Prytulak, L. S. (1974). Sometimes frequently means seldom: Context effects in the interpretation of quantitative expressions. *Journal of Research in Personality, 8,* 95–101.

Rohrmann, B. (1978). Empirische Studien zur Entwicklung von Antwortskalen für die sozialwissenschaftliche Forschung. *Zeitschrift für Sozialpsychologie, 9,* 222–245.

Singpurwalla, N. D., & Booker, J. M. (2004). Membership functions and probability measures of fuzzy sets. *Journal of the American Statistical Association, 467,* 867–877.

Spector, P. E. (1976). Choosing response categories for summated rating scales. *Journal of Applied Psychology, 61,* 374–375.

Teigen, K. H., & Brun, W. (2003). Verbal expressions of uncertainty and probability. In D. Hardman (Ed.), *Thinking: Psychological perspectives on reasoning, judgment and decision making* (pp. 125–145). New York: Wiley.

Wallsten, T. S., Budescu, D. V., Rapoport, A., Zwick, R., & Forsyth, B. (1986). Measuring the vague meaning of probability terms. *Journal of Experimental Psychology. General, 115,* 348–365.

Wallsten, T. S., Budescu, D. V., Zwick, R., & Kemp, S. M. (1993). Preferences and reasons for communicating probabilistic information in numerical or verbal terms. *Bulletin of the Psychonomic Society, 31,* 135–138.

Wyatt, R. C., & Meyers, L. S. (1987). Psychometric properties of four 5-point Likert type response scales. *Educational and Psychological Measurement, 47,* 27–35.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8,* 338–353.

Zimmer, A. (1984). A model for the interpretation of verbal predictions. *International Journal of Man–Machine Studies, 20,* 121–134.