FORS PAPERS
Working paper series

Item nonresponse and fuzzy logic

Georg P. Mueller[1]

[1] Fac. of Economics and Social Sciences, University of Fribourg

FORS Working Paper 01-2019

Lausanne, November 2019

FORS Working Paper series

The FORS Working Paper series presents findings related to survey research, focusing on methodological aspects of survey research or substantive research. Manuscripts submitted are papers that represent work-in-progress. This series is intended to provide an early and relatively fast means of publication prior to further development of the work. A revised version might be requested from the author directly.

Further information on the FORS Working Paper Series can be found on www.forscenter.ch

How to cite this document:
Mueller, Georg P. (2019). Item nonresponse and fuzzy logic. *FORS Working Paper Series, paper 2019-1.* Lausanne: FORS. DOI: 10.24440/FWP-2019-00001

If you want to contact the author: Georg.Mueller_Unifr@bluewin.ch

ISSN 1663-523x (online), DOI: 10.24440/FWP-2019-00001

FORS, c/o University of Lausanne, Géopolis, 1015 Lausanne, Switzerland, E-mail: paperseries@fors.unil.ch

SUMMARY

As compared to the problem of unit nonresponse in standardized interviews, the topic of item nonresponse is less well explored. Hence this article attempts to explain item nonresponse by the fuzziness of the related social situation, which is supposed to lead to a failure in the formation of attitudes about the respective interview-item. The author assumes that interview-items have for the respondents a fuzzy truth, which varies on a continuous scale between 0 = “false“ and 1 = “true“. By this conceptualization, some interview-items may have a fuzzy truth in the middle between the mentioned poles, i.e. at the level 0.5 = “indeterminate“. Consequently, the main hypothesis of the present study postulates for this situation an increased level of item nonresponse. The empirical part of the article tests this hypothesis with data of the European Values Study (EVS 2008) about self-reported religiosity. For this purpose it assumes a correspondence between item response- and fuzzy truth-functions. It turns out, that the main hypothesis of the article holds true for the full sample of the EVS.

Keywords: Item nonresponse, item response function, fuzzy logic, European Values Study, religious beliefs

1. INTRODUCTION AND OVERVIEW

Depending on the European country that is analyzed, up to 15.5%[1] of the national residents do not know, whether they are happy or not with their job. An analogous, somewhat lower figure of 9.7%[2] holds for the ignorance about personal religiosity. Both figures are examples of *item nonresponse* in general population surveys. Item nonresponse is usually defined as an interviewee-behavior, where the respondent either refuses to answer a *particular* question or chooses the predefined response option „Don't know". Thus item nonresponse is different from *unit nonresponse,* where the *whole interview* is for various reasons missing.

The existing literature about nonresponse behavior mainly refers to *unit* nonresponse (Koch and Porst 1998, Schnell 1997, Billiet et al. 2007). Only a minority of research papers focuses on item nonresponse and most of them deal with the prevention and imputation of missing values (de Leeuw et al. 2003, Rubin 1987). With a few exceptions like Huisman (1999), Beatty and Herrmann (2002), Messer et al. (2012), and Krosnick (2002), there seems to be only limited interest in the reasons why certain interviewees do not answer particular questions. Consequently, this article attempts to explore one of the reasons for item nonresponse in standardized interviews about self-attributed religiosity.

One of the possible causes, on which this paper focuses is the *fuzziness* of the social situation to be evaluated by the interviewee. This corresponds to the concept of *ambivalence* of Krosnick (2002: 94-95), coined for interviewees with no opinion about the concerned topic: whereas questions about gender or marital status may be answered by a crisp *yes* or *no,* questions about self-attributed religiosity or satisfaction with a given life-situation often have less determinate answers. The evaluation of a fuzzy situation may be highly volatile, very complex, or contradictory with regard to its different facets. Hence this paper proposes the use of multi-valued fuzzy logic, where the truth of an evaluation or a belief is a continuous variable, which has only two really *crisp* values 1 = true and 0 = false*.* All other truth-values in between are *fuzzy,* the maximum of fuzziness being near 0.5, which points to the total indeterminacy of the situation. Hence, the present article postulates that this kind of indeterminacy is one of the major sources of *nonresponse* to the related interview questions. Consequently, this paper explores the conditions, under which the fuzzy truth of an evaluation or judgment gets close to the value 0.5 of perfect indeterminacy. This implies on the one hand the formulation of appropriate *hypotheses.* On the other hand it requires the development of *procedures for measuring* the fuzzy truth of a judgment.

Both, measurement procedures and hypotheses will be tested by item nonresponse to the question of self-reported *religiosity,* which was included in the European Values Study (EVS 2008a, variable V114). Since the nonresponse rate is generally small, it is important that the EVS 2008 has the advantage of being a very big data-collection with nearly 70'000 respondents from different countries, who were all asked the same questions in their national languages. This makes the international EVS superior to other, national or local data collections. As a matter of course, theory and methodology can also be applied to *single* countries with many cases of item nonresponse: there, the proposed methodology has the advantage of showing the limits of the usual imputation-procedures (Huisman 1999, 96 ff.; Särndal and Lundström 2005, chap. 12), which do not make sense, if missing values are primarily the result of personal undecidedness and not of negligence or lack of willingness when filling the questionnaire.

[1] See variable V90 of Kosovo in EVS (2008a).

[2] See variable V114 of Montenegro in EVS (2008a).

2. ITEM RESPONSE FUNCTIONS AND FUZZY TRUTH [3]

Item response functions are one of the basic tools of psychometric and attitude measurement (Baker and Kim 2004, chap. 1; Hambleton et al. 1991, chap. 2): they describe the mathematical *relation* between an independent variable X such as the frequency of prayers and a certain type of response to an attitude-question, e.g. the belief of being a religious person. This relation generally returns the probability prob(Y = y | X) that for a specific value of the variable X the respondent gives the reply Y = y. The dependent variable Y is often binary with two possible values y = „yes“ and y = „no“. In the case of y = "yes", the mentioned relation is a monotonically increasing function of X, which has often the shape of a logistic- or ramp-function (see Fig. 1a). However, for other response categories like „dk“ = „don‘t know“, it may also have different geometrical forms, like e.g. the peak in Fig. 1c.

Fuzzy truth has been introduced by Zadeh (1965) and others (e.g. Zimmermann 1987) in order to come closer to our everyday reasoning. It is an extension of classical Boolean logic (Bergmann 2008, chap. 2) and thus includes not only the extremes 0 = „false“ and 1 = „true“ but also *all* other values in between (Ragin 2008, chap. 2; Ragin 2000, chap. 6). Thus, fuzzy truth expresses the degree of certainty, to which we hold our beliefs for true. If it refers to a varying property X, it is also possible to construct a *fuzzy truth function*

$$f(Y \mid X) \dashrightarrow [0,1] \quad (1)$$

which returns for the different values of X the fuzzy truth of a given *proposition Y.* A statement Y like „It is unjust to claim more than X percent of the cake“ may thus have a varying truth, depending on the value of the variable X. Truth values near 0.5 are neither true nor false but *indeterminate* in the sense of the three-valued logic of Lukasiewic (1970) (see also Nguyen and Walker 1997, chap. 4.4). Based on other empirical studies like Krosnick (2002, 94-95), we suppose that this indeterminacy of the truth values around 0.5 results in an increased propensity to nonresponse behavior. Hence we postulate

Hypothesis 1:

The greater for a given value of X the *absolute* difference between the fuzzy truth of the item response Y = "yes" and the indeterminate fuzzy truth 0.5, the lower the probability of an item nonresponse Y = "dk" (don't know). Thus, as depicted in Fig. 1c, we are postulating for the different values of X a negative correlation

$$\text{corr}\,[\,|\,f(Y = \text{yes} \mid X) - 0.5\,|\,,\ \text{prob}(Y = \text{dk} \mid X)\,] < 0 \quad (2)$$

By definition, fuzzy truth is *different* from probability (Ross et al. 2002). As mentioned before, the *fuzzy truth* of an event A is the (subjective) certainty that A is true, whereas the *probability* of A is the relative frequency of the occurrence of A. The mathematics of fuzzy truth are also different from the rules of processing probabilities: for example, the *probability*

$$\text{prob}(A \text{ AND } B) = \text{prob}(A) * \text{prob}(B) \quad (3)$$

(see Everitt 2006, 269) of the simultaneous occurrence of two independent stochastic events A and B is different from the *fuzzy truth*

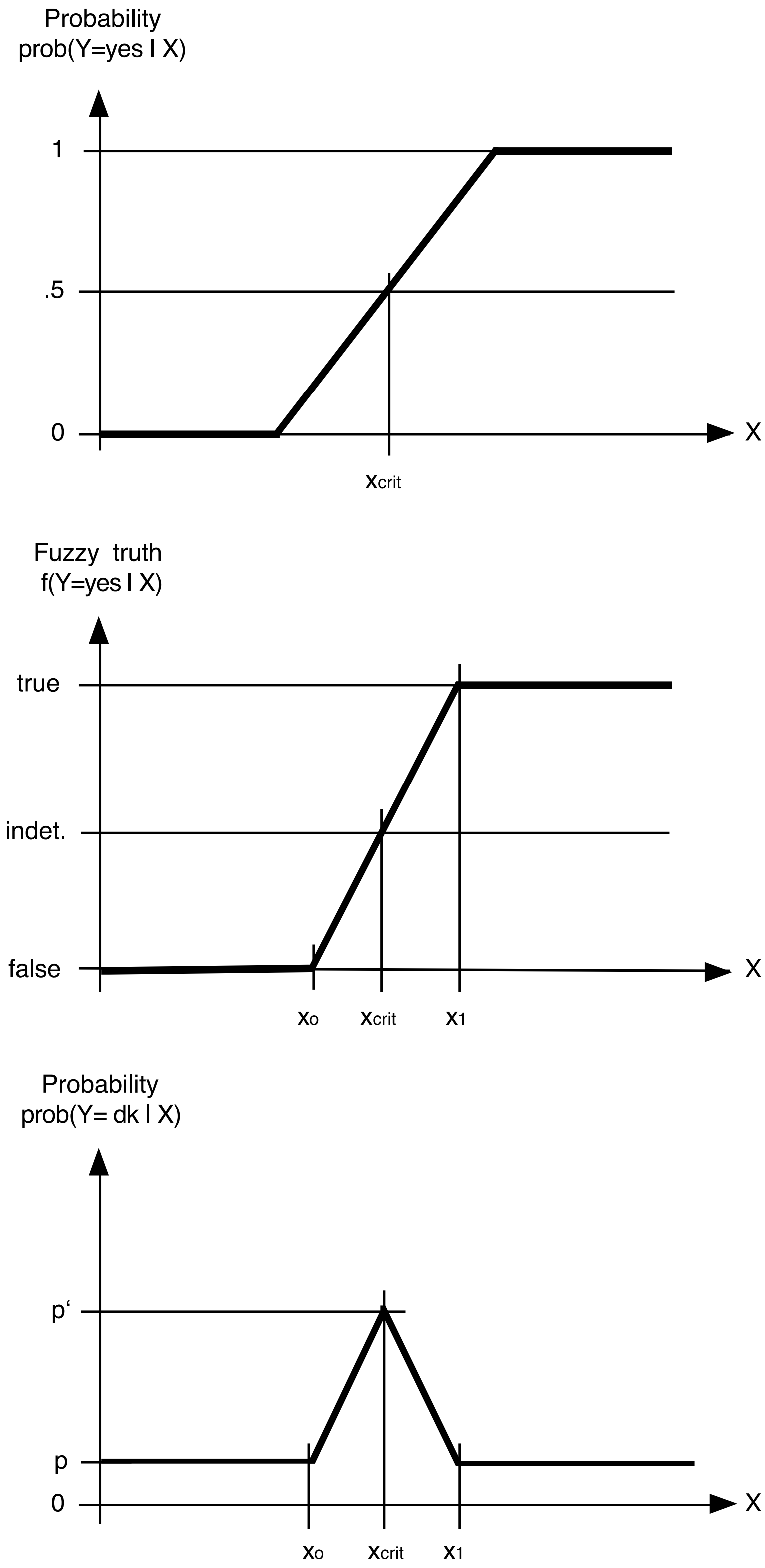
$$f(A \text{ AND } B) = \min(f(A), f(B)) \quad (4)$$

(see Bergmann 2008, chap. 11.2) that A and B have both occurred, even under the assumption that the fuzzy truths f(A) = prob(A) and f(B) = prob(B). Nonetheless, the common probabilistic

[3] For mathematical terms in this and the following sections see Appendix 2 at the end of the article.

aspects of the fuzzy truth- and the item response function suggest the following two ancillary hypotheses (see Figs. 1a,b):

Figure 1a,b,c: The correspondence between item response probability (above), fuzzy truth (middle), and the probability of item nonresponse (below)



Legend: indet. = Indeterminate fuzzy truth; dk = Don‘t know;
x_{crit} = Critical value of X, where the fuzzy truth of Y = yes is indet.;
x_o , x_1 = Start and end of the ramp of the fuzzy truth function;
p = Basic, „technical“ level of missings; p‘ = Max. level of missings

Hypothesis 2a:

The fuzzy truth f(Y= yes | X) and the probability prob(Y= yes | X) are positively correlated:

corr [f(Y= yes | X) , prob(Y= yes | X)] > 0 (5)

Hypothesis 2b:

For the fuzzy truth f(Y= yes | X) = 0.5 there is a one-to-one correspondence with the response probability prob(Y= yes | X), such that

f(Y= yes | X) = 0.5 ⇔ prob(Y= yes | X) = 0.5 (6)

Hypotheses 2a and 2b are in so far crucial for rest of this article, as they create a duality between item nonresponse probabilities and fuzzy truth, which is important for operationalizing the latter concept. Among others, these hypotheses allow to identify the value of X, for which the *probability of nonresponse* reaches the maximum. According to Figs. 1b and 1c this is the case, if the *fuzzy truth* f(Y= yes | X) = 0.5 = indet. (indeterminate). Hence, due to hypothesis 2b we have to look for the value of X, for which the *probability*

prob(Y= yes | X) = 0.5 (7)

On the grounds of Fig. 1a, equation (7) holds true for the value of X = x_{crit}.

3. THE AGGREGATION OF FUZZY TRUTH FUNCTIONS

The item response Y = "yes" often depends not only on one, but on *several* independent variables or components X, X', X", ... As a consequence there is, like in multivariate regression, the problem of evaluating their *joint* effect on the dependent variable Y. In fuzzy logic this requires the integration of the different single *component* truth-functions f(X), f'(X'), f''(X"), ... into a truth function F(X, X', X", ...), which yields the *aggregate* fuzzy truth of Y = "yes". In principle F(X, X', X", ...) can be any logical expression, which combines the truth-functions f(X), f'(X'), f''(X"), ... by one or several of the following fuzzy truth operators:

a) The operator OR, defined by

   F(X,X') = f(X) OR f'(X') = max(f(X) , f'(X')) (8)

   (Bergmann 2008, chap. 11.2). Here the fuzzy truth F(X,X') depends on the largest component of the expression. Consequently, *one* fully true component is sufficient in order to get F(X,X') = 1, independently of the truth of the other component. This obviously corresponds to the functioning of the OR-operation in "classical" binary logic.

b) The operator AND, defined by

   F(X,X') = f(X) AND f'(X') = min(f(X) , f'(X')) (9)

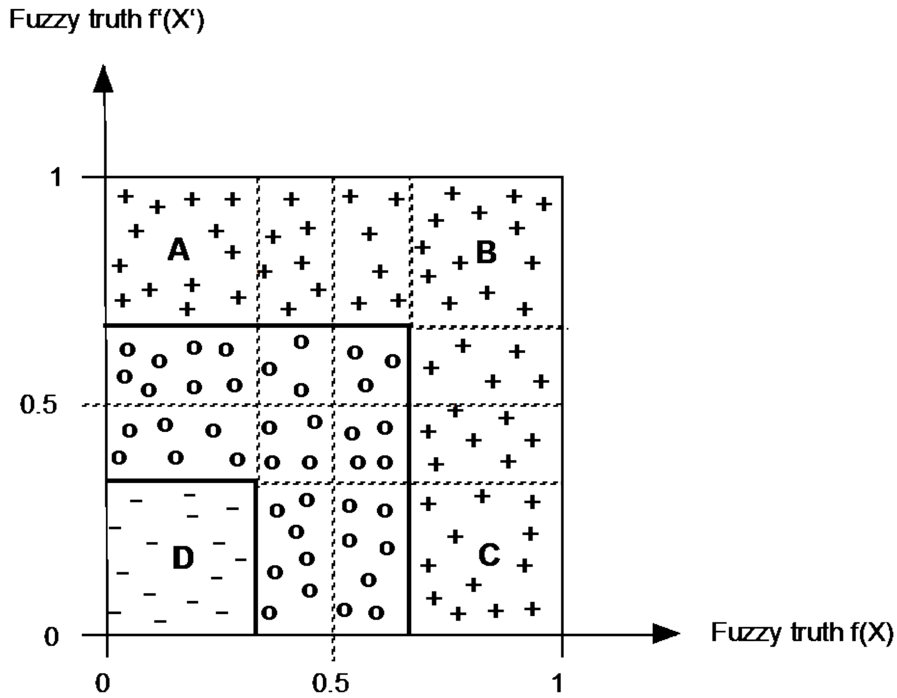
   (see Bergmann 2008, chap. 11.2). Hence the fuzzy truth of F(X,X') depends on the smaller of its two components. Thus, in order to get "full" truth F(X,X') = 1 all components have to be fully true, like in "classical" binary logic.

c) The operator NOT, defined by

   F(X) = NOT f(X) = 1- f(X) (10)

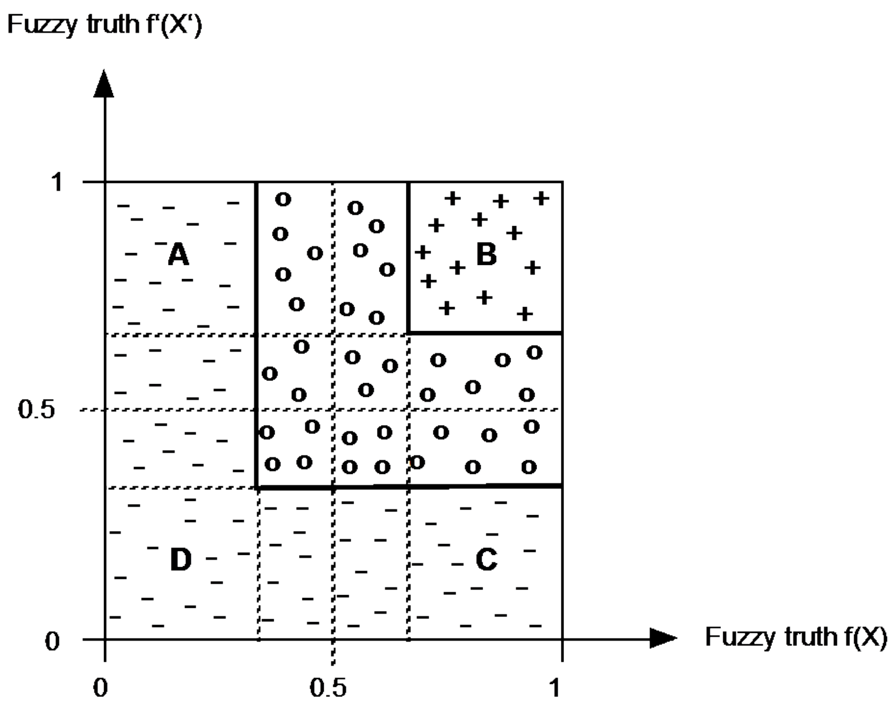
   (Bergmann 2008, chap. 11.2). Hence in fuzzy logic the expression NOT f(X) OR f(X) is generally not "fully" true (= 1), but equals only the bigger of the values 1 - f(X) and f(X). This illustrates one of the differences between fuzzy and "classical" binary logic.

Figure 2a: The aggregated fuzzy truth F(X,X') = f(X) OR f'(X')



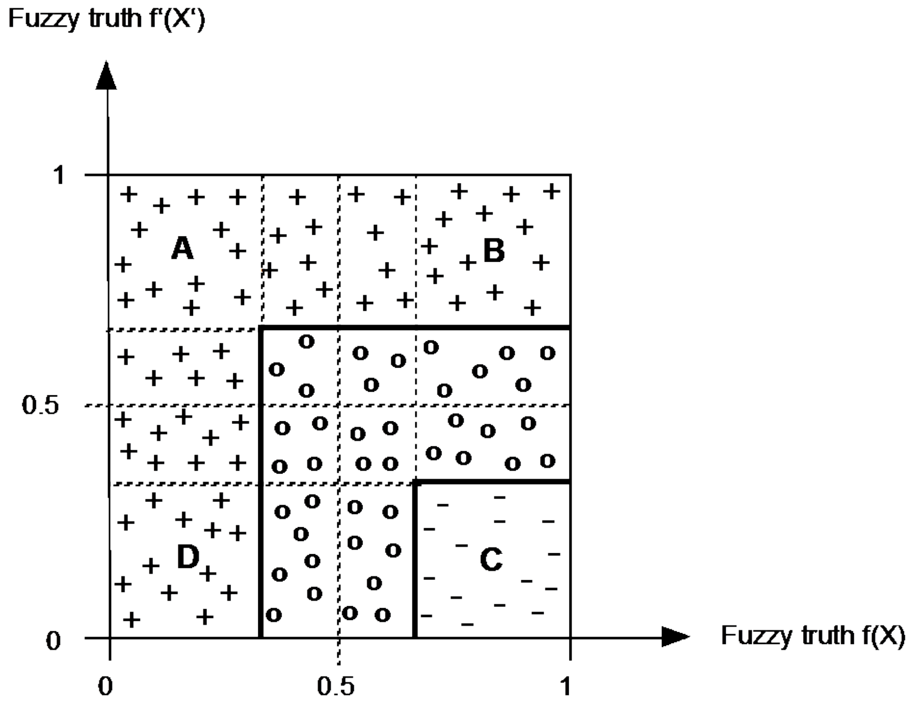
Legend: Aggregated truth F(X,X') = f(X) OR f'(X') = max(f(X), f'(X')).
o: F(X , X') ≈ .50; +: F(X , X') >> .50; -: F(X , X') << .50

Figure 2b: The aggregated fuzzy truth F(X,X') = f(X) AND f'(X')



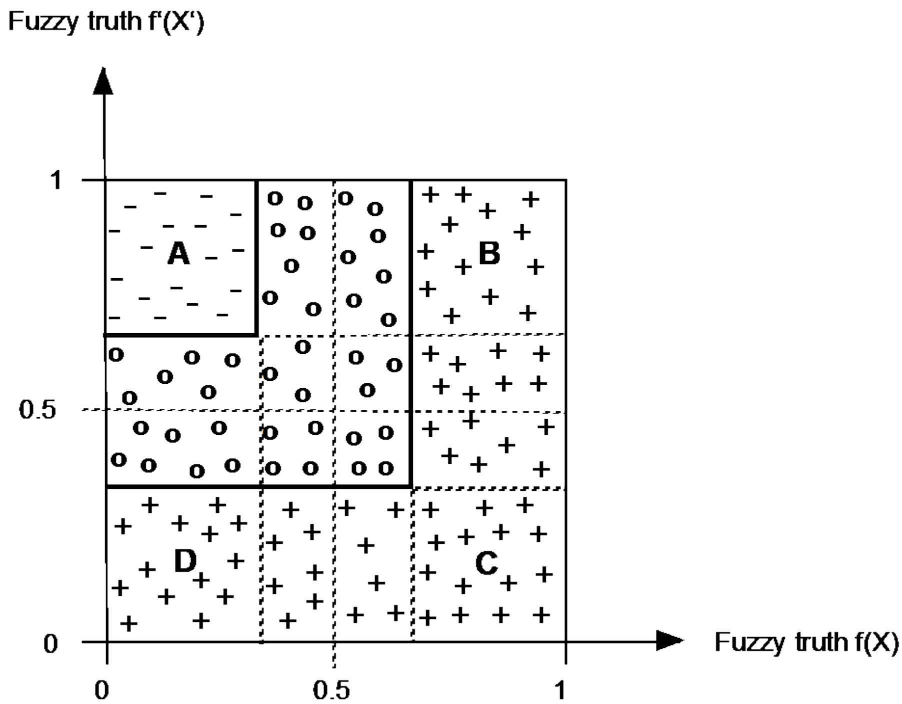
Legend: Aggregated truth F(X,X') = f(X) AND f'(X') = min(f(X), f'(X')).
o: F(X , X') ≈ .50; +: F(X , X') >> .50; -: F(X , X') << .50

Figure 2c: The aggregated fuzzy truth F(X,X') = f(X) ⇨ f'(X')



Legend: Aggregated truth F(X,X') = f(X) => f'(X') = max(1 - f(X), f(X')).
o: F(X , X') ≈ .50; +: F(X , X') >> .50; -: F(X , X') << .50

Figure 2d: The aggregated fuzzy truth F(X,X') = f'(X') ⇨ f(X)



Legend: Aggregated truth F(X,X') = f'(X') => f(X) = max(1- f'(X'), f(X)).
o: F(X , X') ≈ .50; +: F(X , X') >> .50; -: F(X , X') << .50

d) The operator ⇨, i.e. the implication of Kleene-Dienes (Grabowski 2017), defined by

$$F(X,X') = f(X) \Rightarrow f'(X') = \max(\ 1\text{-}f(X)\ ,\ f'(X')\ ) \tag{11}$$

Since max(1-f(X),f'(X')) = NOT f(X) OR f'(X') (see equations (8) and (10)) this definition has the advantage over others that it corresponds to the implication of binary logic (Spies 2004, Tab. 2.1).

The decision about the best possible aggregation of fuzzy components may be theory- and/or data-driven. In both cases the multidimensional distribution of the fuzzy truths may be helpful, either for testing or for developing hypotheses about the aggregation of fuzzy truth-functions.

Fig. 2a illustrates the situation of the fuzzy *OR-aggregation,* e.g. describing the case, when frequent prayers, high importance of God, or both are required for true religiosity: in area D of Fig. 2a f(X) and f'(X') are both relatively low and consequently the aggregated fuzzy truth F(X,X') = max(f(X),f'(X')) is also rather low (code -). In zones A and B of Fig. 2a the fuzzy truth f'(X') is comparatively high and consequently F(X,X') = max(f(X),f'(X')) is too rather high (code +). A similar reasoning explains, why high values f(X) entail high aggregated truth values F(X,X') in the areas B and C (code +). If two adjacent zones A, B, C, or D have *different signs,* the area in between is *coded as 0,* pointing to an aggregated truth value F(X,X') ≈0.5 = indeterminate.

Fig. 2b describes the distribution of the truth values of the *AND-aggregation:* F(X,X') is only in zone B rather high (code +), since

$$F(X,X') = f(X)\ \text{AND}\ f'(X') = \min(\ f(X)\ ,\ f'(X')\ ) \tag{9}$$

can only be high if *both* components f(X) and f'(X') are high. For the remaining zones A, C, and D this is *not* the case and consequently the aggregated truth F(X,X') is low (code -). If e.g. frequent prayers and high importance of God are *both* required for true religiosity, the absence of even one of these criteria reduces the truth of the claim of being a religious person.

Finally, Figs. 2c,d display the results for the aggregation by a *Kleene-Dienes implication,* i.e. f(X) ⇨ f'(X') (Fig. 2c) or f'(X') ⇨ f(X) (Fig. 2d). Here, three out of the four areas A, B, C, and D have a *high* aggregate fuzzy truth (code +). There are however one zone C in Fig. 2c and one zone A in Fig. 2d, which both represent a *low* aggregated truth (code -): in these zones a condition with a high level of fuzzy truth implies an outcome with a low component truth, which means a *wrong* implication. E.g., for religious persons a high importance of God should imply frequent prayers. If it does not, it reduces the truth of the self-attributed claim of being religious to a relatively *low* level.

As suggested earlier, the four Figs. 2a,b,c,d represent typical situations which often occur, when fuzzy components are aggregated by a logical standard operator OR, AND, or ⇨. They are obviously not exhaustive. In principle, there are 4 diagrams, where *one* of the areas A, B, C, or D has a *low* aggregated truth (code -), there are 4 other diagrams, where *one* of these areas has a *high* aggregated truth (code +), and there are 6 diagrams where *two* of the zones A, B, C, or D have *low* aggregated truths (code -). In the previous paragraphs, we presented just 4 out of these 4 + 4 + 6 = 14 possible configurations.

4. EMPIRICAL TESTS

4.1 "BEING A RELIGIOUS PERSON" AS A FUZZY CONCEPT

In section 4 we are going to test the main *hypothesis 1* of this article (see equation (2)) by means of the item nonresponse to the self-evaluation of *being a religious person.* The related variable *Religiosity* has typical properties of a *fuzzy concept:*

Table 1: The fuzziness of concepts, by rates of item nonresponse

Question:	EVS-variable	% item non-response	Total N
Having the local nationality	V304	0.2%	67786
Having a job	V89	0.4%	67786
Being a religious person	V114	4.5%	67286

Legend: Source: EVS (2008a); Sample: All interviewed Europeans, independent of their nationality; Total N: Sample size corresponding to 100% of the analyzed cases.

a) There is *no objective measurement* that allows to rank persons on a commonly agreed scale. Consequently, contradictory claims of superiority with regard to a fuzzy concept cannot be settled in the same way as e.g. in the case of age, salary, or educational attainment. Thus, *subjectivity* prevails and it is easily possible that two respondents both believe that they are more religious than the other.

b) For fuzzy concepts the overall probability of *item nonresponse* should be clearly *higher than zero,* due to a relatively high proportion of cases with indeterminate truth, which according to Fig. 1c increase the item nonresponse rate. As compared to other exemplary interview questions about *having the local nationality* or *having a job,* the concept of *being a religious person* is indeed fuzzy with increased rates of item nonresponse, as shown in Tab. 1.

c) Fuzzy concepts are often *multidimensional* and the different components do not necessarily lead to the same assessment of the situation. As a consequence, rank orders of groups of persons may vary with the aspect of the fuzzy concept which is in focus. This is typically happens to the concept of religiosity, which has multiple facets, such as the belief in the importance of God, the frequency of church-going, the frequency of private prayers, the belief in heaven and/or hell, etc. For the subsequent tests we are using the *Importance of God* and the *Frequency of prayers.* The first variable is certainly a core indicator of religiosity with high prima facie validity and the second may be considered as the factual proof of this belief. Thus we postulate an ancillary *hypothesis 3* about the aggregated fuzzy truth of religiosity:

   $$F\,(\text{Religiosity}) = f\,(\text{Importance of God}) \Rightarrow f'\,(\text{Frequent prayers}) \quad (12)$$

   As a consequence, in order to make the claim of high religiosity true, a high *Importance of God* should imply a high *Frequency of prayers.* If it does not, the aggregated fuzzy truth F(Religiosity) is reduced (see area C in Fig. 2c).

4.2 THE OPERATIONALIZATION OF THE MAJOR CONCEPTS

Item nonresponse is also for fuzzy concepts a rather *rare* event. Consequently, a conventional one-country survey with 1000-2000 interviews is only exceptionally, like in the case of *Sweden,* large enough for a valid empirical study about item nonresponse. A possible solution to this problem of scarce data is the aggregation of several surveys, which consequently multiplies the number of available observations with item nonresponse. This approach, however, requires the *comparability* of interview questions of different countries. Hence it is suggested to use international surveys, where identical questions have been asked in many different countries. Thus we rely in this article on the European Values Study EVS (2008a), which has the advantage over similar projects like the International Social Survey Programme (ISSP 2013) that it is focused on a culturally homogeneous context, i.e. Europe. In particular, we used from this survey the following variables as basic sources of information for our analyses:

a) Religiosity = "yes" for the self-evaluation of being a *religious person*, corresponding to the EVS-variable V114. Religiosity = "no" for codes *not religious* or *convinced atheist* of the EVS-variable V114. Religiosity = "item nonresponse" for the codes "don't know" or "no answer" of the original EVS-variable V114. For further details see Tab. 8 in appendix 1.

b) 8-V132 = *Frequency of prayers* (apart from religious services). The polarity of the original EVS-variable V132 was reversed such that a higher score reflects a higher frequency of prayers. (Details in Tab. 8).

c) V129 = *Importance of God* (for the respondent's personal life). (Details in Tab. 8).

According to the ancillary hypotheses 2a,b there should be a positive correlation

$$\text{corr}\,[\,f(Y= \text{yes} \mid X)\,,\ \text{prob}(Y= \text{yes} \mid X)\,] > 0 \qquad (5)$$

and an equivalence

$$f(Y= \text{yes} \mid X) = 0.5 \Leftrightarrow \text{prob}(Y= \text{yes} \mid X) = 0.5 \qquad (6)$$

between the fuzzy truth f(Y= yes | X) and the item-response function prob(Y= yes | X), where X is an explanatory variable of Y. Consequently, for determining the fuzzy truth functions f(Importance of God) and f'(Frequent prayers) we made two separate *binary logistic regressions* (Hambleton et al. 1991, 12 ff.) in order to explain *Religiosity* by the *Frequency of prayers* on the one hand and by the *Importance of God* on the other. Both independent variables are assumed to have a *positive* effect on *Religiosity.* The resulting item response function

$$\text{prob(Religiosity = yes | Importance of God)} = 1 / \left(1 + e^{-(c + b \,*\, \text{Importance of God})}\right) \qquad (13)$$

is subsequently used as a proxy for the truth function f(Importance of God) and consequently should fulfill hypotheses 2a,b. Similarly,

$$\text{prob(Religiosity = yes | Frequency of prayers)} = 1 / \left(1 + e^{-(c' + b' \,*\, \text{Frequency of prayers})}\right) \qquad (14)$$

is used as an estimate of the truth function f'(Frequent prayers).

4.3 A FIRST TEST WITH SWEDISH DATA

Sweden is one of the few countries, which used for the EVS 2008 *mailed* questionnaires (EVS 2008b). Probably due to this particular method of data collection, Sweden has a nearly sufficient number of missing values for the self-evaluation of being a religious person. Thus we dared to test our hypotheses with Swedish data. In order to do so, we started with the determination of the fuzzy truth functions f(Importance of God) and f'(Frequent prayers). The results of the related binary regressions are displayed in Tab. 2 and show high values for Nagelkerke's r^2 (= Nagelk. r^2) and the percentage of correct binary predictions (= % correct). As expected in section 4.2, both coefficients b are positive and statistically highly significant. Thus we use the related logistic equations (13) and (14) as proxies for the component truth functions f(Importance of God) and f'(Frequent prayers).

For the aggregation of the mentioned fuzzy truth components we proposed in ancillary hypothesis 3

F (Religiosity) = f (Importance of God) ⇨ f' (Frequent prayers) (12)

Table 2: Logistic regressions for estimating the effects of the independent variables Importance of God and Frequency of prayers on Religiosity: The case of Sweden

Independent variable:	Const. c	Coeff. b	Nagelk. r^2	% correct	N of obs.
Importance of God	-3.316 [<0.001]	0.605 [<0.001]	0.595	84.4%	1077
Freq. of prayers	-2.894 [<0.001]	0.849 [<0.001]	0.518	82.2 %	1077

Legend: Independent variables: Importance of God = variable V129 of EVS (2008a); Freq. of prayers = 8 - V132 of EVS (2008a). Dependent variable: Religiosity = yes = Self-evaluation as a religious person. [ ]: One-tailed error probabilities.

If this hypothesis were correct, Tab. 3 should correspond to Fig. 2c. This is obviously *not* the case, because in zone C of Tab. 3 more than 50% of the Swedish consider themselves as religious persons, whereas Fig. 2c postulates that the aggregated fuzzy truth and the related share of religious persons should be much lower than 0.50. The diagram among the Figs. 2a-d that fits best to the observed distribution of religious persons is Fig. 2a, which postulates an OR-aggregation and has like Fig. 3 the *minimal* share of religious people in zone D. Thus we have to revise hypothesis 3 as follows (see eqn. (8) and Tab. 2):

F (Religiosity) = f (Importance of God) OR f' (Frequent prayers) =

= max(f(Importance of God) , f'(Frequent prayers)) =

$$= \max(1 / (1 + e^{-(-3.316 + 0.605 * \text{Importance of God})}) , 1 / (1 + e^{-(-2.894 + 0.849 * \text{Frequency of prayers})})) \quad (15)$$

Hence, it seems that religiosity is in Lutheran Sweden rather abstract and does not require a factual proof in the form of frequent prayers.

Table 3: % of Swedish with self-evaluation as religious persons, by configuration of the fuzzy truths

		f (Importance of God) < 0.40	f (Importance of God) > 0.60
f' (Frequent prayers)	> 0.60	A: 75.0 % [12]	B: 90.5 % [190]
f' (Frequent prayers)	< 0.40	D: 7.3 % [599]	C: 57.9 % [38]

Legend: A,B,C,D: Configurations of fuzzy truth, as described in Figs. 2a-d. [ ]: N of observations of a configuration, excluding missing values. f(Importance of God): Fuzzy truth of religiosity by importance of God. f'(Frequent prayers): Fuzzy truth of religiosity by frequency of prayers.

Table 4: % of persons with nonresponse about their religiosity, by aggregated fuzzy truth F(Religiosity): The case of Sweden

Intervals of aggregated fuzzy truth F(Religiosity)														
< .25		.25 –.35		.35 –.45		*.45 –.55*		.55 –.65		.65 –.75		. 75 –.85		≥.85
6.2%	**<**	14.9%	**>**	3.9%	**<**	*19.0%*	>	15.4%	>	5.8%	>	4.9%	<	5.1%
[529]		[174]		[51]		*[42]*		[39]		[52]		[41]		[236]

Legend: Aggregated fuzzy truth F(Religiosity) = f (Importance of God) OR f' (Frequent prayers). Intervals of F(Religiosity): Excluding upper limits. [ ]: N corresponding to 100%. Bold **>** , **<** : Order of % of nonresponse is significant at α ≤5% (1-tailed z-test). Underlined: Values consistent with theory.

By means of equation (15) we are now able to test our basic hypothesis 1: The further away the aggregated fuzzy truth F(Religiosity) is from the value of indeterminacy 0.5, the lower the probability prob(Religiosity = dk) of an item nonresponse (see Fig. 1c). Tab. 4 seems to confirm this hypothesis, at least to a certain extent: the maximum rate of item nonresponse is – as theoretically expected – at F(Religiosity) = 0.50 +/- 0.05, with a considerable level of 19.0% missings. For higher aggregated truth values F(Religiosity) > 0.50, the share of nonresponse systematically *decreases,* with one negligible exception: F(Religiosity) ≥0.85. Similarly, for lower aggregated truth values F(Religiosity) < 0.50, the share of item nonresponse is also systematically decreasing, however with the exception of F(Religiosity) = 0.40 +/- 0.05 (see Tab. 4). Moreover, many theoretically expected changes in Tab. 4 are statistically not significant. Thus, contrary to our original expectations the sample size of Sweden is probably not large enough for a statistically perfect confirmation of hypothesis 1.

4.4 A SECOND TEST WITH EUROPEAN DATA

In view of the unsatisfactory data of Sweden, we are adding a second test that is based on an unweighted pooling of all 46 countries, which participated in the EVS 2008. As before in the case of Sweden, we started with the estimation of the fuzzy truth components f(Importance of God) and f'(Frequent prayers). The results of the binary logistic regressions described in section 4.2 are presented in Tab. 5. As expected in section 4.2 both coefficients b are positive and statistically

Table 5: Logistic regressions for estimating the effects of the independent variables Importance of God and Frequency of prayers on Religiosity: The case of Europe

Independent variable:	Const. c	Coeff. b	Nagelk. r^2	% correct	N of obs.
Importance of God	-2.294 [<0.001]	0.581 [<0.001]	0.536	86.0%	62645
Freq. of prayers	-1.445 [<0.001]	0.728 [<0.001]	0.461	84.1 %	62238

Legend: Independent variables: Importance of God = variable V129 of EVS (2008a); Freq. of prayers = 8 - V132 of EVS (2008a). Dependent variable: Religiosity = yes = Self-evaluation as a religious person. [ ]: One-tailed error probabilities.

highly significant. Together with the high Nagelkerke r^2 (= Nagelk. r^2) and the high percentage of correct binary predictions (% correct) Tab. 5 gives us confidence that the results of the logistic regressions may be used for further analyses.

As a first subsequent step we calculated for different fuzzy truth configurations of f(Importance of God) and f'(Frequent prayers) the share of the religious persons. The results are presented in Tab. 6 and *contradict* the ancillary hypothesis 3, which postulates

F (Religiosity) = f (Importance of God) ⇨ f' (Frequent prayers) (12)

By comparing Tab. 6 with Figs. 2a and 2c the OR-aggregation is much more likely than the postulated implication ⇨: Fig. 2a and Tab. 6 both have the smallest shares of religious persons in area D, while in the remaining zones A, B, and C the respective percentages are substantially greater than 50%. Thus frequent prayers and the importance of God are each an independent source of religiosity. At least for the catholic and orthodox this is plausible, since religious persons address prayers also to saints (e.g. St-Mary) and not only to God.

Table 6: % of Europeans with self-evaluation as religious persons, by configuration of the fuzzy truths

		f (Importance of God)	
		< 0.40	> 0.60
f' (Frequent prayers)	> 0.60	A: 65.7 % [1434]	B: 92.6 % [36706]
f' (Frequent prayers)	< 0.40	D: 10.5 % [10954]	C: 65.6 % [3680]

Legend: A,B,C,D: Configurations of fuzzy truth, as described in Figs. 2a-d. [ ]: N of observations of a configuration, excluding missing values. f(Importance of God): Fuzzy truth of religiosity by importance of God. f'(Frequent prayers): Fuzzy truth of religiosity by frequency of prayers.

Table 7: % of persons with nonresponse about their religiosity, by aggregated fuzzy truth F(Religiosity): The case of Europe [4]

Intervals of aggregated fuzzy truth F(Religiosity)														
< .35		.35 –.45		*.45 –.55*		.55 –.65		.65 –.75		.75 –.85		. 85 –.95		≥.95
5.8%	**<**	8.7%	<	*9.2%*	>	8.7%	>	7.9%	**>**	5.8%	**>**	3.3%	**>**	1.6%
[10918]		[1682]		*[3539]*		[3533]		[1885]		[4169]		[15568]		[25380]

Legend: Aggregated fuzzy truth F(Religiosity) = f (Importance of God) OR f' (Frequent prayers). Intervals of F(Religiosity): Excluding upper limits. [ ]: N corresponding to 100%. Bold **>** , **<** : Order of % of nonresponse is significant at α ≤5% (1-tailed z-test). Underlined: Values consistent with theory.

[4] The value 5.8% for F(Religiosity) < .35 includes a relatively small number of 160 cases (= 1.5% of N = 10918), which belong to the category F(Religiosity)=[.15,.25[and have a deviant nonresponse rate of 17.5%.

Due to the empirical results of Tabs. 5 and 6, we continue our analyses on the basis of the assumption that

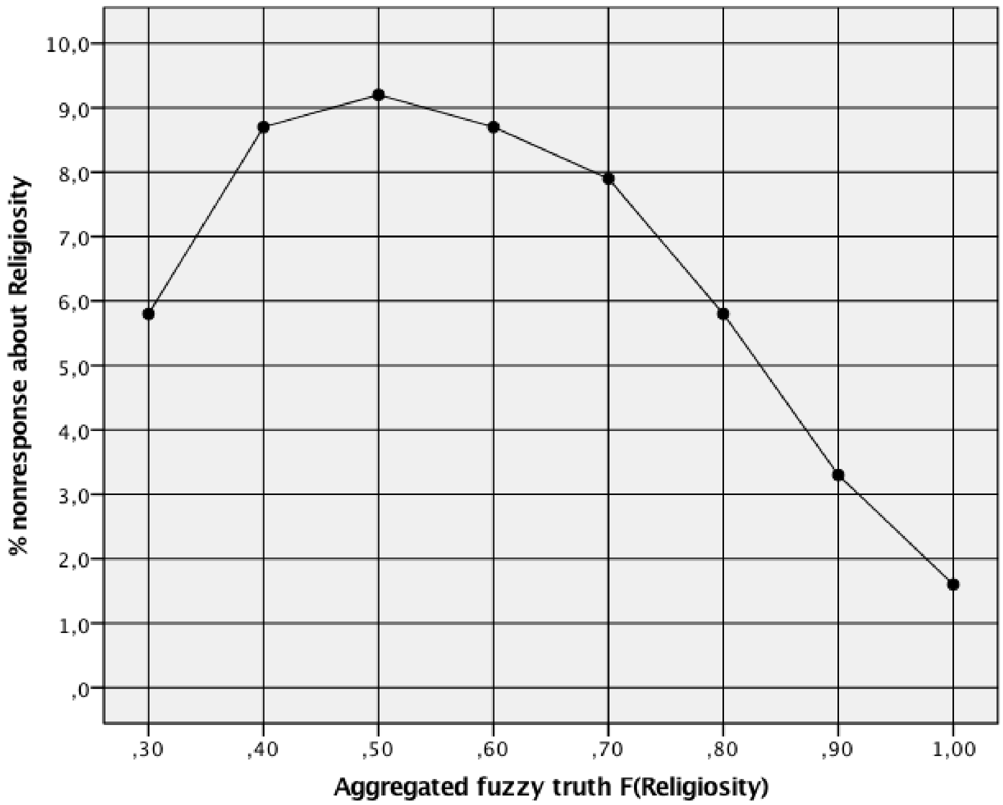
F (Religiosity) = f (Importance of God) OR f' (Frequent prayers) =

= max(f (Importance of God) , f'(Frequent prayers)) =

$$= \max(1 / (1 + e^{-(-2.294 + 0.581 * \text{Importance of God})}) , 1 / (1 + e^{-(-1.445 + 0.728 * \text{Frequency of prayers})})) \quad (16)$$

This allows us to investigate the relation between the aggregated fuzzy truth of equation (16) and the share of the item nonresponse to the question about religiosity. According to hypothesis 1 and Fig. 1c we expect an inverse u-relation with a maximum of item nonresponse at F(Religiosity) ≈ 0.50. Tab. 7 and Fig. 3 confirm this theoretical expectation, although the differences between adjacent entries in Tab. 7 are not always statistically significant. As a consequence of the inverse u-relation of Fig. 3, item nonresponse with regard to *Religiosity* is unlikely, if the *Importance of God* and the *Frequency of prayers* are both either very low or very high: these situations are for the interviewed persons crisp and well structured. However, if one of the mentioned factors is at medium level and the other low or medium, the risk of item nonresponse with regard to *Religiosity* rises according to equation (16) to its maximum. Since item nonresponse is in this situation mainly the result of personal undecidedness and less of negligence or lack of willingness when filling the questionnaire, the usual imputation-procedures (Huisman 1999, 96 ff.; Särndal and Lundström 2005, chap. 12) should be handled with care: they pretend a hidden opinion that does not really exist.

Figure 3: The relation between the aggregated fuzzy truth F(Religiosity) and the % of nonresponse about Religiosity: The case of Europe



Legend: Fuzzy truth F(Religiosity): The lowest category includes all cases with F(Religiosity) < .35; the highest category includes all values F(Religiosity) ≥.95.
Source: Tab. 7.

5. SUMMARY AND OUTLOOK

In this article we intended to show the usefulness of fuzzy logic for analyzing and explaining item nonresponse. We postulated in hypothesis 1 that interview answers with a fuzzy truth around the value 0.5 have an increased probability of item nonresponse. The further away the fuzzy truth is from this critical value 0.5, the lower the probability of nonresponse.

In order to test this hypothesis we first postulated a correspondence between item response- and fuzzy truth-functions. This way fuzzy truth became measurable by means of binary logistic regressions. Since there are often not only one but several item response- and related fuzzy truth-functions, we also had to solve the problem of aggregating the different component truth functions. The natural solution to this problem was the use of the common operators of fuzzy logic, i.e. AND, OR, NOT, and the implication ⇨.

With this methodological toolbox it was possible to test the main hypothesis 1 by investigating the cases of missing answers to the self-evaluation of being a religious person. Our empirical analyses confirmed our main hypothesis 1 that aggregated fuzzy truth around the value 0.5 increases the rate of nonresponse, however with two reservations: *First,* the aggregation of the fuzzy components had to be based on the OR-operator instead of the implication ⇨, which is only of minor importance for the test of the main hypothesis 1. Yet, such a revision of the original assumptions illustrates an important feature of the present methodology: it allows to identify the interaction between the explanatory variables of the analyzed item nonresponse. *Second,* the *changes* of the nonresponse rates between adjacent intervals of aggregated fuzzy truth are not always statistically significant, although the directions of these changes are generally correct. Thus hypothesis 1 should perhaps be re-tested with other fuzzy concepts, which have more cases of item nonresponse than religiosity and consequently entail less problems with statistical beta-errors.

REFERENCES

Baker, F. & Kim, S.-H. (2004). *Item Response Theory.* Boca Raton: CRC Press.

Beatty, P. & Herrmann, D. (2002). To Answer or Not to Answer: Decision Processes Related to Survey Item Nonresponse. In R. Groves et al. (Eds.) *Survey Nonresponse,* chap. 5. New York: Wiley.

Bergmann, M. (2008). *An Introduction to Many-Valued and Fuzzy Logic.* Cambridge: Cambridge University Press.

Billiet, J., Koch, A. & Philippens, M. (2007). Understanding and Improving Response Rates. In R. Jowell et al. (Eds.) *Measuring Attitudes Cross-Nationally,* chap. 6. Los Angeles: Sage.

De Leeuw, E., Hox, J. & Huisman, M. (2003). Prevention and Treatment of Item Nonresponse. *Journal of Official Statistics 19,* 153–176.

Everitt, B. S. (2006). *The Cambridge Dictionary of Statistics.* Cambridge: Cambridge University Press.

EVS (2008a). *European Values Study 2008.* http://www.gesis.org/unser-angebot/daten-analysieren/umfragedaten/european-values-study/4th-wave-2008/

EVS (2008b). *European Values Study 2008: Sweden.* https://dbk.gesis.org/DBKSearch/SDESC2.asp?no=4761&DB=E

Grabowski, A. (2017). Formal Introduction to Fuzzy Implications. *Formalized Mathematics 25,* 241–248.

Hambleton, R., Swaminathan, H. & Rogers, J. (1991). *Fundamentals of Item Response Theory.* Newbury Park: Sage.

Huisman, M. (1999). *Item Nonresponse: Occurrence, Causes, and Imputation of Missing Answers to Test Items.* Leiden: DSWO Press.

ISSP (2013). *International Social Survey Programme.* http://www.issp.org/

Koch, A. & Porst, R. (Eds.) (1988). *Nonresponse in Survey Research.* Mannheim: Zuma.

Krosnick, J. (2002). The Causes of No-Opinion Responses to Attitude Measures in Surveys: They Are Rarely What They Appear to Be. In R. Groves et al. (Eds.) *Survey Nonresponse,* chap. 6. New York: Wiley.

Lukasiewicz, J. (1970). *Selected Works* (ed. by L. Borkowski). Amsterdam: North-Holland.

Messer, B. L., Edwards, M. L., & Dillman, D. A. (2012). Determinants of Item Nonresponse to Web and Mail Respondents in Three Address-Based Mixed-Mode Surveys of the General Public. *Survey Practice*, *5*(2), 1-8.

Nguyen, H. & Walker, E. (1997). *A First Course in Fuzzy Logic.* Boca Raton: CRC Press.

Ragin, Ch. (2000). *Fuzzy-Set Social Science.* Chicago: University of Chicago Press.

Ragin, Ch. (2008). *Redesigning Social Inquiry: Fuzzy Sets and Beyond.* Chicago: University of Chicago Press.

Ross, T., Sellers, K., & Booker, J. (2002). Considerations for Using Fuzzy Set Theory and Probability Theory. In T. Ross, K. Sellers & J. Parkinson (Eds.) *Fuzzy Logic and Probability Applications: Bridging the Gap,* chap. 5. Philadelphia: ASA-SIAM Press.

Rubin, D. (1987). *Multiple Imputation for Nonresponse in Surveys.* New York: Wiley.

Särndal, C.-E. & Lundström, S. (2005). *Estimation in Surveys with Nonresponse.* Chichester: Wiley.

Schnell, R. (1997). *Nonresponse in Bevölkerungsumfragen (Nonresponse in Population Surveys).* Opladen: Leske + Budrich.

Spies, M. (2004). *Einführung in die Logik (Introduction to Logic).* Heidelberg: Spektrum Akademischer Verlag.

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

Zimmermann, H. (1987). *Fuzzy Sets, Decision Making, and Expert Systems.* Boston: Kluwer.

APPENDIX 1 (TABLE 8): DESCRIPTIVE STATISTICS OF THE MAIN VARIABLES OF THE EMPIRICAL ANALYSES

Variable:		Sweden:		Europe:	
	Value:	Cases:	Percent:	Cases:	Percent:
Religiosity	not asked	0	0.0%	500	0.7%
	no answer	14	1.2%	851	1.3%
	don't know	96	8.1%	2153	3.2%
	religious	365	30.7 %	45927	67.8%
	not religious	563	47.4%	14981	22.1%
	atheist	149	12.6%	3374	5.0%
Importance of God	no answer	9	0.8%	1065	1.6%
	don't know	85	7.2%	1178	1.7%
	not at all	420	35.4%	8889	13.1%
	2	123	10.4%	2904	4.3%
	3	98	8.3%	3154	4.7%
	4	40	3.4%	2513	3.7%
	5	81	6.8%	6575	9.7%
	6	44	3.7%	4350	6.4%
	7	56	4.7%	5533	8.2%
	8	57	4.8%	7227	10.7%
	9	31	2.6%	5117	7.5%
	very	143	12.0%	19281	28.4%
Frequency of prayers	no answer	12	1.0%	1192	1.8%
	don't know	33	2.8%	1428	2.1%
	never	626	52.7%	17105	25.2%
	2	198	16.7%	7256	10.7%
	3	80	6.7%	5090	7.5%
	4	35	2.9%	4681	6.9%
	5	15	1.3%	4897	7.2%
	6	61	5.1%	7806	11.5%
	every day	127	10.7%	18331	27.0%

Legend: Religiosity = EVS variable V114 "Are you a religious person?". Importance of God = EVS variable V129 "How important is God in your life?". Frequency of prayers = 8 - EVS variable V132 "How often do you pray to God outside religious services?"

APPENDIX 2: GLOSSARY OF MATHEMATICAL TERMS

A *AND* B:	Logical AND-operator. In *binary* logic the result is only true, if the propositions A and B are both true. In *fuzzy* logic the result is equal to the lower of the fuzzy truth values of A and B.
Coeff. b :	Unstandardized regression coefficient b.
Const. c :	Unstandardized regression constant c.
corr (X,Y):	Spearman correlation between X and Y.
e :	Euler's number (= Basis of the natural logarithm).
f(X), f'(X'), ..:	Fuzzy truth functions of the components X, X', ...
F(X, X', ...) :	Aggregate fuzzy truth function of the components X, X', ...
indet.	Indeterminate fuzzy truth value 0.5 .
max (X, Y):	The bigger of the values X and Y.
min (X, Y):	The smaller of the values X and Y.
N of obs.:	Number of observations.
Nagelk r^2 :	Nagelkerke's r^2 as a measure of the goodness of fit.
NOT A:	Logical negation. In *binary* logic the truth of the result is the reverse of the truth of the proposition A. In *fuzzy* logic the result is 1 minus the fuzzy truth of A.
A *OR* B:	Logical OR-operator. In *binary* logic the result is only true, if A or B or both are true. In *fuzzy* logic the result is equal to the higher of the fuzzy truth values of A and B.
prob (A):	Probability of A.
% correct :	Percentage of correct binary predictions of a logistic regression.
A ⇨ B:	Logical implication. In *binary* logic the result is only true if the propositions A and B are both true or A is false. In *fuzzy* logic the result is equal to the higher of the fuzzy truth values of NOT A and B.
$\vert X \vert$:	Absolute value of X.
X ≈Y:	Approximate equality of X and Y.
X << Y:	X is substantially smaller than Y.
X >> Y:	X is substantially bigger than Y.