International Journal of Approximate Reasoning ••• (••••) •••–•••

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Typicality: A formal concept analysis account

Radim Belohlavek, Tomas Mikula

Department of Computer Science, Palacky University, Olomouc, 17. listopadu 12, CZ-771 46 Olomouc, Czech Republic

ARTICLE INFO

Article history:
Received 28 February 2021
Received in revised form 25 October 2021
Accepted 1 December 2021
Available online xxxx

Keywords:
Typicality
Categorization
Concept
Formal concept analysis
Psychology of concepts

ABSTRACT

We examine typicality—a significant phenomenon accompanying human concepts—within the framework of formal concept analysis. Our aim is to formalize the notion of typicality within this framework and thus provide an operational definition. We review relevant aspects and main psychological explanations of typicality, and propose a formalization based on a view of typicality propounded in the seminal work by Eleanor Rosch et al. We also provide experimental evaluation of our approach and discuss ramifications of our findings and topics to be explored in the future.

1. Aims of our study

Concepts are at the center of human reasoning and are hence the subject of numerous explorations. Among them, psychological studies of concepts have a distinguished role. The psychology of concepts provides a number of interesting theories and experimental studies of various phenomena involving concepts. These are of interest not only for the domain of psychology itself but, naturally, also for other domains concerned with concepts, including numerous formal approaches to reasoning and information processing using concepts.

One of the most significant phenomena accompanying human concepts, which is broadly familiar from everyday life, is typicality: A sparrow is a typical bird, an ostrich is not; a trout is a typical fish, an eel or a flounder is not.[1] Typicality may be regarded as a manifestation of a graded structure of concepts, and plays a remarkable role in several cognitive tasks, such as categorization and classification, which are crucially important in processing information by humans.

Our aim in this paper is to formalize a view of typicality propounded in the seminal works on typicality and the graded structure of concepts by Eleanor Rosch et al. For our purpose, we utilize the framework of formal concept analysis (FCA).[2] This framework is naturally suited for our purpose because its fundamental notions, such as that of object, attribute, sharing of attributes, as well as other notions, appear as basic in most of the psychological studies of typicality. In a sense, selection of FCA represents a choice of a straightforward simple framework that provides formal counterparts to the primitive notions used informally by psychologists.

Two particular motivations for our study are as follows. First, formalization of typicality allows us to approach and explore typicality in precise terms amenable to formal analysis. This is important particularly in view of the fact that in the psychological literature, theories of typicality are described rather informally, very often just verbally. Similarly informal are reasoning and the conclusions regarding typicality presented in the literature. Formalization of typicality, on the other hand,

E-mail addresses: radim.belohlavek@acm.org (R. Belohlavek), mail@tomasmikula.cz (T. Mikula).

[1] In our cultural context; the role of cultural context is mentioned below.

[2] We thus continue our previous effort to examine the basic level of concepts [5–7] within FCA.

https://doi.org/10.1016/j.ijar.2021.12.001
0888-613X/

renders an operational definition which lets one realize the various subtleties and possible shortcomings of an informal, verbally described definition of typicality. In addition, it enables one to consider possible relationships to alternative definitions and related notions, and thus may generally help examine typicality in a more rigorous manner.[3]

Secondly, we believe that formalization of typicality is significant for FCA itself. In general, we consider extensions of data analytical and information processing methods, such as FCA, by notions coming from the psychology of concepts a meaningful task which may significantly enhance these methods. While typicality has as yet not been exploited in FCA, it seems a natural mean to extend the structure of formal concepts.

Our paper is meant to make first steps in studying typicality in the framework of FCA. In section 2, we provide an overview of typicality from the viewpoint of the psychology of concepts and present selected issues pertaining to typicality. Our formalization of typicality within the FCA framework is outlined in section 3. Examples and experiments involving typicality are the subject of section 4. Conclusions and a prospect of further topics to explore is outlined in section 5.

2. Psychological accounts of typicality

2.1. Typicality as manifestation of a graded structure of concepts

Until the 1970s, the prevalent paradigm in psychological studies of concepts was represented by the so-called classical view.[4] According to this view, a concept is determined by a set of yes/no (bivalent, binary) conditions (attributes, features) which are necessary and jointly sufficient, i.e. definitory in the following sense: An object is covered by (or, is a member of) the concept (or category in terms commonly used in the psychology of concepts) if and only if the object satisfies each of these conditions. This view has a long tradition in philosophy and logic and also underlies the notion of a concept in the basic setting of formal concept analysis.[5]

In the mid-1970s, various explorations—most importantly those led by Eleanor Rosch—in the internal structure of concepts revealed fundamental limitations of the classical view. Put briefly, it became apparent that concepts have a graded structure: Various phenomena had experimentally been found to be a matter of degree rather than bivalent (yes/no). In addition, important phenomena had been observed that were not accounted for by the classical view. Typicality, which is discussed in the first findings by Rosch et al. [22–24], represents such a phenomenon. Ever since these first findings, the phenomenon continues to be a subject of vivid psychological research; see e.g. [10,27].

The classical view does not account for typicality, at least not directly, which represents a shortcoming of this view. Namely, according to the classical view, all members of a category have an equal status with respect to the category. On the other hand, people naturally regard some objects more typical of a given category than other objects. Further research has shown that people are even capable of assigning degrees of typicality (called also typicality ratings) to objects for a given category in a consistent manner.

Note in this connection that another phenomenon, which had been examined in the early 1970s, that involves degrees and is not addressed by the classical view is the graded nature of a membership in category itself. That is, an object may not just be a member or a non-member of a given category, but rather a member to a certain degree in the sense of fuzzy sets.[6] While the classical view is constrained to two possible degrees of membership, namely 0 (non-member) and 1 (member), the more general view, which is experimentally confirmed as significantly more appropriate, allows for degrees of membership, such as 0.8 representing high but not full membership or 0.5 representing a borderline case.

Basically, there are two possible views to start from in considering typicality. The literature on the psychology of concepts does not, unfortunately, make it clear to which of these views a particular study of typicality subscribes; see. e.g. [20]. In the first view, membership in a category is bivalent (i.e. classical, yes/no) and typicality represents an additional structure of a category. In the second view, membership is graded and possibly even equivalent to (or otherwise strongly correlated with) typicality. In our formalization below we assume the former view, i.e. that categories (concepts) are classical and that typicality represents an additional structure. Such view is adopted, e.g., in the design of experiments in the seminal paper [23].

Note also an important feature of typicality, namely its high cognitive significance; see e.g. [1,20,23]. For one, people tend to agree on typicality ratings. Moreover, typicality is reported to predict performance in a variety of cognitive tasks including learning of categories (typical objects are learned more quickly), deciding membership in categories (decisions on typical objects are more quick), and production of category exemplars (typical exemplars are generated first). Typical items are also useful in making inferences about categories and serve as so-called cognitive reference points.

[3] This aspect was a significant part of our work on basic level [5–7].

[4] A detailed exposition of developments in the psychological theories of concepts is provided in the monograph [20]; see also [18].

[5] Note, however, that in the fuzzy logic extension of the basic setting of formal concept analysis, the attributes are considered fuzzy (graded) rather than bivalent; see e.g. [2,3].

[6] Note that Rosch's studies of graded nature of categories were conducted independently and in about the same time as Zadeh's studies of fuzzy sets [28]. Both Rosch and Zadeh were with UC Berkeley.

2.2. Explanations of typicality

In their seminal paper [23], Rosch and Mervis put forward a hypothesis of what makes an object typical in a category. This hypothesis was confirmed by experiments by the authors [23] and had later been examined by numerous other studies; see e.g. the monograph [20], in which typicality occupies a significant part. Rosch and Mervis [23, p. 575] describe their hypothesis as follows:

> ...members of a category come to be viewed as prototypical of the category as a whole in proportion to the extent to which they bear a family resemblance to (have attributes that overlap those of) other members of the category. Conversely, items viewed as most prototypical of one category will be those with least family resemblance to or membership in other categories.

The first part referring to resemblance (similarity) to objects of the given concept (category) is intuitively compelling and relatively straightforward to formalize. It is this part that we use in our approach. The second part referring to resemblance to objects in other concepts is not so straightforward, brings non-trivial problems, which are also reflected in the experiments in [23], and we hence do not consider it in what follows.[7]

In addition, several other possible explanations of typicality of an item have been suggested and tested in later studies, including similarity to central tendency (central tendency being e.g. the average of a numerical characteristic of an item), closeness to ideals in goal-oriented categories (ideals represent characteristics that items should possess if they are to serve the goal associated with a category), frequency of instantiation (i.e. frequency of encounter with the item as a member of a given category), and familiarity (i.e. frequency of encounter across all contexts); see e.g. [1,19,20] and also [14]. A more recent research emphasizes also the role of context (situation) in which typicality is assessed [27]. The resulting instability of typicality resulting from dependence on context even led the authors in [10] to distinguish between the so-called structural typicality (representing stability) and functional typicality (representing context-dependence and thus instability).

In spite of several alternative hypotheses, the family resemblance hypothesis of Rosch and Mervis [23] mentioned above appears to remain the most simple and most commonly accepted explanation of typicality. It is due to this fact that Rosch and Mervis' explanation forms the basis of our approach to typicality within the framework of FCA.

3. Formalization of typicality within formal concept analysis

3.1. Preliminaries from formal concept analysis (FCA)

FCA [9,15] starts with its basic notion of a formal context, which is a triplet $\langle X, Y, I\rangle$ consisting of non-empty sets X and Y, and a binary relation (incidence relation) I between X and Y (that is, $I \subseteq X \times Y$, i.e. I consists of selected object-attribute pairs $\langle x, y\rangle$). The sets X and Y are interpreted as the set of objects and the set of (yes/no) attributes, and the fact $\langle x, y\rangle \in I$ means that the object x has the attribute y. An example of a formal context is depicted in Table 5 in section 4: Objects $x \in X$ and attributes $y \in Y$ are represented by table rows and columns, and the incidence relation I by crosses and blanks; for $x =$ scorpion and $y =$ predator we have $\langle x, y\rangle \in I$ (scorpion is predator), for $x =$ frog and $y =$ hair we have $\langle x, y\rangle \notin I$ (frog does not have hair), etc.

A pair $\langle A, B\rangle$ consisting of a set $A \subseteq X$ of objects and a set $B \subseteq Y$ of attributes is called a formal concept in $\langle X, Y, I\rangle$ if and only if $A^{\uparrow} = B$ and $B^{\downarrow} = A$ where

$$A^{\uparrow} = \{y \in Y \mid \text{for each } x \in A : \langle x, y\rangle \in I\},$$
$$B^{\downarrow} = \{x \in X \mid \text{for each } y \in B : \langle x, y\rangle \in I\}.$$

Notice that $A^{\uparrow}$ and $B^{\downarrow}$ are the set of all attributes common to all objects in A and the set of all objects having all the attributes in B, respectively. Geometrically, formal concepts in $\langle X, Y, I\rangle$ are maximal rectangular areas (up to a permutation of rows and columns) in the table representing $\langle X, Y, I\rangle$ that are full of crosses. The notion of a formal concept corresponds to the traditional notion of concept as consisting of its extent (objects covered by the concept) and its intent (attributes characterizing the concept); the extent of a formal concept $\langle A, B\rangle$ is A, the intent is B.

In a given formal context $\langle X, Y, I\rangle$, there is, as a rule, a number of formal concepts. The set of all formal concepts in a given formal context $\langle X, Y, I\rangle$ is denoted by $\mathcal{B}(X, Y, I)$, i.e.

$$\mathcal{B}(X, Y, I) = \{\langle A, B\rangle \mid A^{\uparrow} = B \text{ and } B^{\downarrow} = A\},$$

and is called the concept lattice of $\langle X, Y, I\rangle$. Namely, when equipped with a natural subconcept-superconcept hierarchy $\le$, defined by

[7] The problem is with the meaning of "other categories". We leave this problem for future research. Note, however, that the properties mentioned in the first part (i.e. similarity to objects of the given category, which we use) and the second part (small similarity to objects in other categories) were tested separately in [23], and that each of these two parts was found significantly correlated with typicality ratings.

$$\langle A, B\rangle \leq \langle C, D\rangle \text{ if and only if } A \subseteq C \text{ (equivalently, if and only if } B \supseteq D),$$

the set $\mathcal{B}(X, Y, I)$ indeed becomes a complete lattice, whose structure is described by the basic theorem of concept lattices [15].

3.2. Our approach to typicality

As noted above, psychological explorations of typicality and other facets of the graded structure of concepts are considered a strong argument against the classical view of concepts. Since FCA is rooted in the classical view of concepts, one might conclude that using FCA is not appropriate for modeling typicality. In our view, this is not the case. We contend that typicality naturally occurs even in concepts with a yes/no membership resulting from yes/no attributes, as in the basic setting of FCA. Moreover, the seminal psychological experiments on typicality mentioned above, as well as several other studies of typicality in the psychological literature are based on the idea of objects described by yes/no attributes.

Let $\langle A, B\rangle \in \mathcal{B}(X, Y, I)$ be a formal concept in a given formal context $\langle X, Y, I\rangle$. In accordance with Rosch and Mervis' view of typicality (section 2.2), we intend to regard an object x as typical for the given concept $\langle A, B\rangle$ to the extent to which it is similar to the objects in A, i.e. to the objects of this concept. A straightforward way is to assume a function

$$sim : X \times X \to [0, 1] \tag{1}$$

assigning to every two objects $x_1, x_2 \in X$ a number $sim(x_1, x_2) \in [0, 1]$ that may be interpreted as a degree to which x_1 and x_2 are similar (we come back to these functions below). Similarity of x to the objects x_1 in A, which underlies Rosch and Mervis' view of typicality, may naturally be interpreted as the average similarity of x to all the objects $x_1 \in A$. This leads to the following definition[8]:

Definition 1. Given a similarity (1), a *degree of typicality* of object $x \in A$ in a formal concept $\langle A, B\rangle \in \mathcal{B}(X, Y, I)$ with $A \neq \emptyset$ is defined by

$$typ(x, \langle A, B\rangle) = \frac{\sum_{x_1 \in A} sim(x, x_1)}{|A|}. \tag{2}$$

Remark 1. (a) Admittedly, our approach is restrictive. One might, for instance, consider formula (2) for x not necessarily in A, or consider the notion of typicality of a subconcept, rather than an object, in a given concept. We proceed with our definition for simplicity.

(b) Typicality degrees provide additional information about a concept $\langle A, B\rangle$. Namely, they reveal a certain graded structure of the concept $\langle A, B\rangle$. Such a structure has a cognitive significance and may be further utilized. Notice that since $typ(x, \langle A, B\rangle) \in [0, 1]$ due to $sim(X, X) \subseteq [0, 1]$, the mapping $t : A \to [0, 1]$ defined by $t(x) = typ(x, \langle A, B\rangle)$ may be regarded as a fuzzy set [28] of objects typical of $\langle A, B\rangle$.

(c) The idea of an element being similar to other elements in a given set has been explored in the literature on clustering and machine learning in general; see e.g. [17,29] on typicality in clustering, and the literature on medoids in clustering, e.g. [21], and silhouettes in clustering, e.g. [25].

Let us now consider the choice of the similarity function (1). It seems natural to derive the degree $sim(x_1, x_2)$ to which the objects x_1 and x_2 are similar from the descriptions of these objects in terms of attributes, i.e. from the sets $\{x_1\}^{\uparrow}$ and $\{x_2\}^{\uparrow}$ (note that $\{x\}^{\uparrow}$ is the set of attributes possessed by x). We hence assume that

$$sim(x_1, x_2) = sim_Y(\{x_1\}^{\uparrow}, \{x_2\}^{\uparrow}), \tag{3}$$

where

$$sim_Y : 2^Y \times 2^Y \to [0, 1]$$

is a function assigning to arbitrary subsets B_1 and B_2 of the set Y of given attributes a degree $sim_Y(B_1, B_2) \in [0, 1]$ that may be interpreted as a degree of similarity of B_1 and B_2. Such functions have been studied in various areas, most notably in the field of clustering; see e.g. [13].

Two particular functions serving this purpose, which we use in our experiments, are the well-known Jaccard index [16], sim_J, and the simple matching coefficient, sim_{SMC}, defined by

[8] Average similarity is mentioned in some psychological studies; see e.g. [1, p. 630]. Note that we also explored minimum instead of average, as it represents the best lower similarity-threshold. Average, nevertheless, yielded more intuitive results. We use $[0, 1]$ for the range (i.e. similarity is scaled), but $\mathbb{R}^+$ is also a natural option (non-scaled).

$$sim_{\mathrm{J}}(B_1, B_2) = \frac{|B_1 \cap B_2|}{|B_1 \cup B_2|} \quad \text{and} \tag{4}$$

$$sim_{\mathrm{SMC}}(B_1, B_2) = \frac{|B_1 \cap B_2| + |Y - (B_1 \cup B_2)|}{|Y|}, \tag{5}$$

respectively. That is, $sim_{\mathrm{J}}(B_1, B_2)$ is the number of attributes that belong to both B_1 and B_2 divided by the number of all attributes that belong to B_1 or B_2; $sim_{\mathrm{SMC}}(B_1, B_2)$ is the number of attributes on which B_1 and B_2 agree (either $y \in B_1$ and $y \in B_2$, or $y \notin B_1$ and $y \notin B_2$) divided by the number of all attributes. Hence, while sim_{SMC} treats both presence and non-presence of attributes symmetrically, sim_{J} disregards non-presence. This is the main conceptual difference between sim_{J} and sim_{SMC}.

The choice of the similarity sim_Y is in a sense crucial and, obviously, several other options different from sim_{J} and sim_{SMC} are possible. In this study, we nevertheless refrain from exploiting the variety of possible further similarity functions. Note, however, that in the next section, we naturally come to a third similarity, which we consider in this paper.

3.3. Relationship to Rosch and Mervis's formula for typicality

Formula (2) for typicality derives in a straightforward (and–as we contend–the most direct) way from the verbal description of Rosch and Mervis's hypothesis quoted in section 2.2. Interestingly, in their experiments to test the hypothesis, Rosch and Mervis [23] use a different formula for typicality of an object. Strangely, this formula does not bear a direct connection to similarity of objects, which is crucial in the hypothesis. The formula is described in [23] as follows. Given a concept, one assigns to every attribute its weight, namely the number of all objects of the concept that possess the attribute. A typicality of a given object in the concept is then the sum of the weights of all the attributes possessed by the object.

This definition translates to the FCA framework as follows. For a given $\langle A, B\rangle \in \mathcal{B}(X, Y, I)$ and $y \in Y$, put

$$w(y, \langle A, B\rangle) = |\{x \in A \mid x \in \{y\}^{\downarrow}\}| \quad \text{(weight of attribute } y\text{)}.$$

Now, according to Rosch and Mervis, the typicality of the object $x \in A$ with respect to the concept $\langle A, B\rangle$ is defined by

$$typ_{\mathrm{RM}}(x, \langle A, B\rangle) = \textstyle\sum_{y \in \{x\}^{\uparrow}} w(y, \langle A, B\rangle).$$

The following theorem shows that in fact, Rosch and Mervis's formula for typicality, which is on the first sight of a different sort compared to our (2), may actually be regarded as resulting from a particular case of our scheme (2) by a simple scaling.

Theorem 1. *For the function* $sim_{\mathrm{rm}}(x_1, x_2) = \frac{|\{x_1\}^{\uparrow} \cap \{x_2\}^{\uparrow}|}{|Y|}$ *we have*

$$typ_{\mathrm{RM}}(x, \langle A, B\rangle) = |A| \cdot |Y| \cdot typ_{\mathrm{rm}}(x, \langle A, B\rangle)$$

where $typ_{\mathrm{rm}}(x, \langle A, B\rangle)$ *is determined by* sim_{rm} *according to (2).*

Proof. Since

$$|A| \cdot |Y| \cdot typ_{\mathrm{rm}}(x, \langle A, B\rangle) = |A| \cdot |Y| \cdot \frac{\sum_{x_1 \in A} sim_{\mathrm{rm}}(x, x_1)}{|A|}$$

$$= |A| \cdot |Y| \cdot \frac{\sum_{x_1 \in A} \frac{|\{x\}^{\uparrow} \cap \{x_1\}^{\uparrow}|}{|Y|}}{|A|} = \sum_{x_1 \in A} |\{x\}^{\uparrow} \cap \{x_1\}^{\uparrow}|,$$

we clearly need to verify

$$typ_{\mathrm{RM}}(x, \langle A, B\rangle) = \sum_{x_1 \in A} |\{x\}^{\uparrow} \cap \{x_1\}^{\uparrow}|.$$

Denoting $||\varphi||$ the truth value of φ (e.g. $||y \in \{x_1\}^{\uparrow}|| = 1$ iff $y \in \{x_1\}^{\uparrow}$), we obtain

$$\sum_{x_1 \in A} |\{x\}^{\uparrow} \cap \{x_1\}^{\uparrow}| = \sum_{x_1 \in A} \sum_{y \in \{x\}^{\uparrow}} ||y \in \{x_1\}^{\uparrow}||$$

$$= \sum_{x_1 \in A} \sum_{y \in \{x\}^{\uparrow}} ||x_1 \in \{y\}^{\downarrow}|| = \sum_{y \in \{x\}^{\uparrow}} \sum_{x_1 \in A} ||x_1 \in \{y\}^{\downarrow}||$$

$$= \sum_{y \in \{x\}^{\uparrow}} |\{x_1 \in A \mid x_1 \in \{y\}^{\downarrow}\}| = \sum_{y \in \{x\}^{\uparrow}} w(y, \langle A, B\rangle) = typ_{\mathrm{RM}}(x, \langle A, B\rangle),$$

completing the proof. □

Remark 2. (a) For a given formal concept $\langle A, B\rangle$ there thus exists a constant c (namely, $c = |A| \cdot |Y|$) such that Rosch-Mervis typicality typ_{RM} is obtained as a c-multiple of a particular typicality (namely typ_{rm}) obtained from our scheme.

(b) As a consequence, the list of objects sorted by typ_{RM} coincides with the list sorted by typ_{rm}.

4. Experiments

We performed experiments with data reported in [11] (section 4.1) and the well-known Zoo data [12] (section 4.2). Our main goal was to observe whether our formulas for typicality agree with human judgment, that is observe to what extent the rankings of objects by typicality degrees (or typicality ratings in terms often used in the psychological literature) computed by our formulas agree with the rankings resulting from human judgment, i.e. human assessment of typicality, for the above data. While human judgment data are available for the data in [11], we had to obtain the human judgment data for the Zoo data by our own questionnaire. Secondly, we attempted to analyze relationships between the formulas for typicality, which we provided, by observing agreements of the typicality degrees computed by the formulas. For the purpose of observing such agreements, both between our formulas and human judgment and between pairs of our formulas, we utilized various rank order correlation coefficients.

4.1. Experiments with Dutch data

Dutch data and the parts used in our experiments

The data used in this section is presented in [11], a study which provides perhaps the most comprehensive data regarding common human categories and their numerous characteristics, including typicality ratings. We first provide a brief description of the data and describe which part we used; the reader is referred to [11] for details.

The data comprises information on both the so-called natural kind and artifact categories, as these two types of categories are believed to have distinct properties (such as mental representation) by the psychologists. The data includes 16 human categories, each of which is represented by a number of selected exemplars (i.e. objects in the sense of FCA). A set of exemplars for a given category is to be considered an extent of the category (we come later to whether it actually is an extent in the sense of FCA). The categories include 10 natural kind categories[9]: “fruit” (30 exemplars); “vegetables” (30); “professions” (30); “sports” (30); the animal categories “amphibians” (5), “birds” (30), “fish” (23), “insects” (26), “mammals” (30), and “reptiles” (22). In addition, they include 6 artifact categories: “clothing” (29 exemplars), “kitchen utensils” (33), “musical instruments” (27), “tools” (30), “vehicles” (30), and “weapons” (20). The exemplars are selected to be representative of the categories; for instance, the animal categories cover a rather large part of the known animal domain.

The Dutch data also contains information on features (attributes in the sense of FCA). Both objects (exemplars) and attributes (features) were obtained by processes described in [11]. For the obtained objects and attributes, data describing which objects have which attributes was also obtained. Consequently, various matrices (called exemplar by feature applicability matrices by the authors) describing which objects have which attributes were obtained. From the FCA viewpoint, these particular matrices represent particular formal contexts $\langle X, Y, I\rangle$ in a straightforward manner (X and Y are the sets of exemplars and features covered by the matrix and I represents which exemplars have which features).

It is to be noted that two ways of obtaining attributes were used in [11]. Respondents were either asked to list attributes for a given category (these are called category attributes) or for a given exemplar (exemplar attributes). These two kinds of attributes are distinct (for example: category features obtained for the category “fish” overlap with the union of exemplar features obtained for the particular exemplars in this category). As a result, one obtains two versions of the applicability matrices: exemplar by category-feature matrices, and exemplar by exemplar-feature matrices.

Each applicability matrix had been filled out separately by four participants in the study (i.e. the participants were filling out whether objects have attributes). To obtain a single matrix out of these four, we required at least two participants to agree. That is, we defined the corresponding formal context $\langle X, Y, I\rangle$ as follows:

$$\langle x, y\rangle \in I \quad \text{iff at least 2 participants claim that } x \text{ has } y$$

In our study, we utilize four of the formal contexts corresponding to the matrices described in the previous paragraphs. These are described by the following table:

dataset	objects	attributes	density
AnimalCategory	129	225	0.32
AnimalExemplar	129	764	0.13
ArtifactCategory	166	301	0.23
ArtifactExemplar	166	1295	0.09

The table describes the numbers of objects and attributes, and the density of the formal context (e.g., 0.32 means that 32% of the 129×225 entries in the AnimalCategory matrix have value 1).

[9] In this list, we use plural in category names, as the authors do; below, we use singular, i.e. “bird” rather than “birds” to be consistent with our previous writings.

Table 1
Typicality ratings for "bird"; AnimalCategory data.

typ_{SMC} order	typ_{SMC}	typ_{J} order	typ_{J}	typ_{rm} order	typ_{rm}	typ_{HJ} order	typ_{HJ}
woodpecker	0.920	woodpecker	0.783	owl	0.295	sparrow	19.179
blackbird	0.917	blackbird	0.780	parrot	0.295	blackbird	18.821
magpie	0.916	cuckoo	0.777	falcon	0.293	robin	18.321
cuckoo	0.916	magpie	0.774	duck	0.292	dove	18.143
robin	0.913	robin	0.770	dove	0.291	crow	18.107
swallow	0.911	swallow	0.766	sparrow	0.288	seagull	17.964
crow	0.910	crow	0.763	parakeet	0.287	canary	17.893
peacock	0.908	chickadee	0.762	swallow	0.287	magpie	17.893
chickadee	0.908	sparrow	0.759	cuckoo	0.285	swallow	17.857
seagull	0.907	falcon	0.757	chickadee	0.285	parakeet	17.643
falcon	0.905	seagull	0.755	blackbird	0.285	chickadee	17.107
sparrow	0.905	owl	0.751	seagull	0.283	eagle	16.926
pheasant	0.904	peacock	0.746	crow	0.282	woodpecker	16.429
pelican	0.902	dove	0.742	woodpecker	0.282	heron	16.107
heron	0.902	parrot	0.741	rooster	0.282	cuckoo	16.000
owl	0.901	pelican	0.739	robin	0.281	owl	16.000
stork	0.901	pheasant	0.737	canary	0.278	parrot	15.857
dove	0.898	canary	0.736	magpie	0.278	falcon	15.500
canary	0.898	parakeet	0.735	chicken	0.276	stork	15.393
parrot	0.896	stork	0.732	pelican	0.275	vulture	15.143
parakeet	0.895	heron	0.729	eagle	0.274	pheasant	13.714
chicken	0.893	chicken	0.724	vulture	0.268	swan	12.821
turkey	0.893	duck	0.721	turkey	0.267	duck	12.786
duck	0.886	turkey	0.718	peacock	0.267	pelican	12.571
rooster	0.884	rooster	0.711	stork	0.266	peacock	12.286
swan	0.882	eagle	0.700	ostrich	0.266	turkey	11.679
eagle	0.881	swan	0.696	swan	0.266	chicken	11.571
ostrich	0.879	ostrich	0.689	pheasant	0.263	ostrich	11.214
vulture	0.865	vulture	0.669	heron	0.258	rooster	11.071
penguin	0.861	penguin	0.653	penguin	0.257	penguin	8.643

An important question is whether the 16 categories, around which the Dutch data is developed and which are represented by sets of exemplars as described above, actually represent formal concepts; that is, whether the sets of exemplars actually form extents of formal concepts in the formal contexts obtained from the considered applicability matrices.[10] Interestingly, most of the 16 categories indeed represent formal concepts. In particular, this is true of the categories "bird," "fish," and "mammal" which we examine in detail below.[11]

Obtained degrees of typicality

For each of the three concepts mentioned above, we computed the degrees of typicality typ_{SMC}, typ_{J}, and typ_{rm} for all objects in the extent of the concept. We present the results for the AnimalCategory data; for the AnimalExemplar data, our observations are similar.[12] The results are shown in Tables 1 ("bird"), 2 ("fish"), and 3 ("mammal"). In addition to the three computed typicalities, the tables also display the typicality degrees obtained by humans, which are part of the Dutch data and which we denote by typ_{HJ}. Note that we keep the values of typ_{HJ} as they are stored in the Dutch data: They range between 1 and 20 since they are obtained as average degrees assigned by respondents on a twenty-element scale. Each table therefore contains four lists of object-typicality pairs, corresponding to the four kinds of typicality (typ_{SMC}, typ_{J}, typ_{rm}, and typ_{HJ}), and each list is ordered by degrees of typicality. The typicality data displayed in the three tables is also displayed in Figs. 1, 2, and 3.

Thus, for instance, the last two columns of Table 1 display a list of bird exemplars sorted by typicality obtained by human judgment along with the typicality degrees: sparrow is the most typical bird by human judgment, followed by blackbird, robin, etc. On the other hand, penguin, rooster and ostrich are considered the least typical of the available exemplars. Intuitively, this sorted list makes sense. By and large, each of the three other lists of exemplars makes intuitively sense as well. One also observes, for the most part, agreement of each of these three lists, which are obtained by our formulas for computing typicality, with the list obtained by human judgment: The birds typical by human judgment generally tend to be typical according to typ_{SMC}, typ_{J}, and typ_{rm}, and the same may be said of untypical birds.

[10] Note that a given a set A is an extent in a given formal context $\langle X, Y, I\rangle$ iff $A = A^{\uparrow\downarrow}$, i.e. the test is straightforward.

[11] Selection of concepts for a detailed exposition is due to lack of space. Our observations below are representative for what we were able to observe for the other concepts and data.

[12] Recall that AnimalCategory and AnimalExemplar have the same sets of objects but differ in their sets of attributes. While the three categories represent formal concepts in both datasets, the computed typicality degrees are different for the two datasets as a result of the difference in attribute sets, because our formulas for typicality depend on the attribute sets.

Table 2
Typicality ratings for "fish"; AnimalCategory data.

typ_{SMC} order	typ_{SMC}	typ_{J} order	typ_{J}	typ_{rm} order	typ_{rm}	typ_{HJ} order	typ_{HJ}
stickleback	0.927	stickleback	0.813	eel	0.291	goldfish	18.893
plaice	0.925	plaice	0.808	salmon	0.289	salmon	18.393
sardine	0.922	sardine	0.804	pike	0.287	cod	18.107
cod	0.921	cod	0.798	stickleback	0.286	trout	17.893
swordfish	0.920	sole	0.797	sardine	0.285	herring	17.071
sole	0.920	carp	0.795	carp	0.284	pike	16.286
pike	0.920	pike	0.795	plaice	0.284	carp	16.000
carp	0.919	trout	0.789	piranha	0.282	plaice	15.962
trout	0.916	salmon	0.789	herring	0.282	eel	15.679
ray	0.915	swordfish	0.788	flatfish	0.282	sardine	15.536
salmon	0.915	herring	0.785	ray	0.280	piranha	15.321
herring	0.915	flatfish	0.783	trout	0.280	sole	15.143
flatfish	0.915	eel	0.780	sole	0.280	stickleback	14.750
eel	0.911	ray	0.779	cod	0.278	swordfish	14.643
anchovy	0.908	anchovy	0.764	anchovy	0.276	ray	14.500
piranha	0.880	piranha	0.708	swordfish	0.275	flatfish	14.321
goldfish	0.878	goldfish	0.687	goldfish	0.259	shark	13.214
squid	0.868	squid	0.661	squid	0.252	anchovy	13.143
shark	0.847	shark	0.624	shark	0.250	squid	10.679
sperm whale	0.840	sperm whale	0.601	sperm whale	0.235	whale	10.429
dolphin	0.812	dolphin	0.561	dolphin	0.231	sperm whale	9.893
whale	0.804	whale	0.552	whale	0.231	orca	9.857
orca	0.803	orca	0.550	orca	0.230	dolphin	9.179

Table 3
Typicality ratings for "mammal"; AnimalCategory data.

typ_{SMC} order	typ_{SMC}	typ_{J} order	typ_{J}	typ_{rm} order	typ_{rm}	typ_{HJ} order	typ_{HJ}
zebra	0.917	zebra	0.754	dog	0.265	cat	18.536
llama	0.914	kangaroo	0.747	cat	0.262	dog	18.536
kangaroo	0.914	llama	0.741	monkey	0.261	monkey	17.929
dromedary	0.911	dromedary	0.738	horse	0.259	lion	17.679
deer	0.908	deer	0.737	lion	0.258	cow	17.607
donkey	0.907	giraffe	0.726	squirrel	0.255	horse	17.536
giraffe	0.906	donkey	0.725	mouse	0.255	sheep	17.429
bison	0.898	horse	0.722	tiger	0.253	pig	17.179
squirrel	0.898	squirrel	0.718	rabbit	0.253	tiger	17.071
horse	0.898	bison	0.714	deer	0.251	wolf	17.036
fox	0.896	fox	0.708	wolf	0.249	donkey	16.821
cow	0.896	cow	0.705	fox	0.249	rabbit	16.643
sheep	0.894	monkey	0.702	bison	0.248	deer	16.536
beaver	0.891	lion	0.699	beaver	0.247	elephant	16.250
hamster	0.889	beaver	0.697	elephant	0.245	fox	16.250
elephant	0.888	sheep	0.697	kangaroo	0.245	zebra	16.036
monkey	0.888	elephant	0.691	zebra	0.244	giraffe	15.964
lion	0.887	wolf	0.691	hamster	0.241	mouse	15.679
rhinoceros	0.886	hamster	0.689	dromedary	0.241	rhinoceros	15.143
wolf	0.886	cat	0.683	cow	0.241	polar bear	15.143
cat	0.877	rabbit	0.677	donkey	0.239	bison	15.143
pig	0.877	mouse	0.676	giraffe	0.239	kangaroo	14.750
rabbit	0.877	tiger	0.675	llama	0.238	llama	14.643
mouse	0.876	rhinoceros	0.673	hedgehog	0.237	hippopotamus	14.607
tiger	0.876	dog	0.664	sheep	0.237	hamster	14.571
hedgehog	0.875	hedgehog	0.658	polar bear	0.232	squirrel	14.571
hippopotamus	0.874	pig	0.656	hippopotamus	0.227	dromedary	14.429
polar bear	0.874	polar bear	0.654	pig	0.227	beaver	14.000
dog	0.865	hippopotamus	0.651	rhinoceros	0.227	hedgehog	13.179
bat	0.834	bat	0.579	bat	0.225	bat	10.857

General remarks on analyzing typicality data

Note at this point two important aspects that need to be kept in mind in our examination. One concerns human judgment on typicality and is known from the literature: Even though human judgment scores are sometimes called the ground truth, the scores may hardly be regarded as objective. Namely, typicality is subjective to a certain extent as it depends on the experience of the respondent, cultural background and other factors. Secondly, since the computed typicalities rely on the available attributes, the attributes need to describe the objects well: they need to describe the domain of inquiry in a

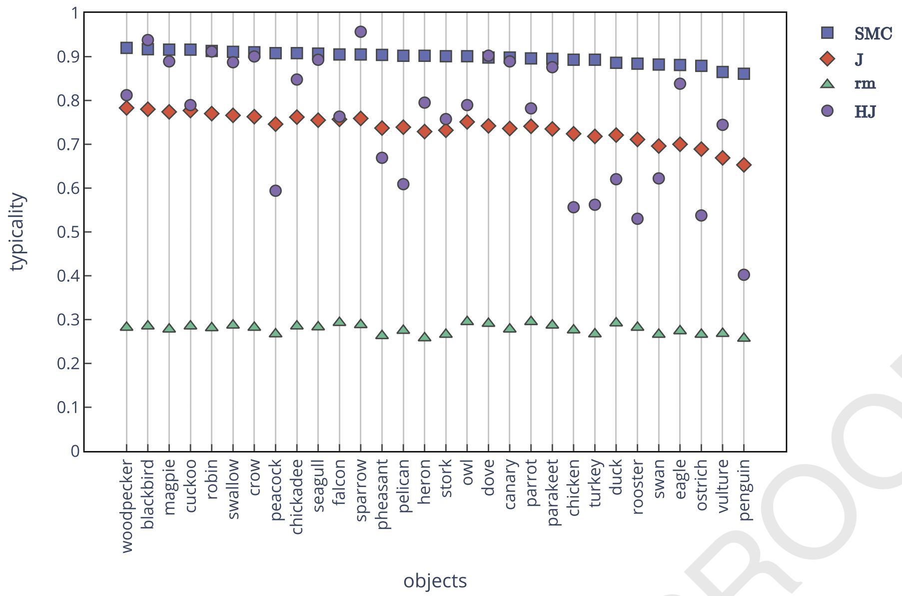

Fig. 1. Typicality ratings for "bird" with objects ordered by values of typ_{SMC} and the values of typ_{HJ} rescaled to $[0, 1]$; AnimalCategory data.

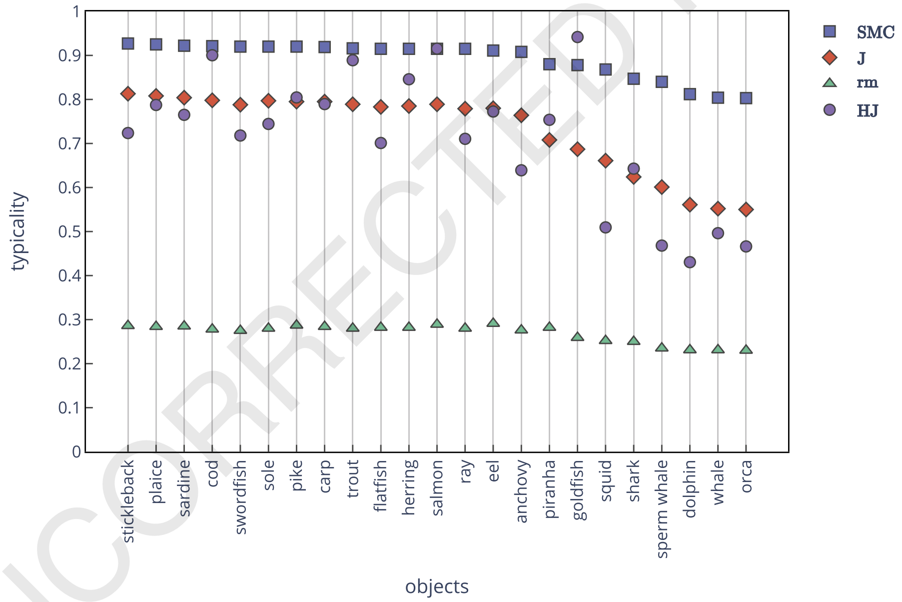

Fig. 2. Typicality ratings for "fish" with objects ordered by values of typ_{SMC} and the values of typ_{HJ} rescaled to $[0, 1]$; AnimalCategory data.

sufficiently informative and balanced way.[13] That is, there need to be enough attributes, describing relevant aspects of the domain, and the attributes must not be redundant (otherwise, the aspect described by redundant attributes would obtain an inappropriately large weight). It is for these reasons that one may hardly expect complete or nearly complete agreement of the computed typicalities with human judgment. Nevertheless, as we demonstrate below, a closer examination reveals reasonable agreements.

Three methods for analyzing typicality data

To assess the agreements of the typicalities and the sorted lists based on them in a more precise manner, we used three methods. First, we used the well-known Kendall tau rank correlation coefficients τ_b. Kendall tau measures ordinal

[13] In the context of examination of basic level, this is also observed in [6,7].

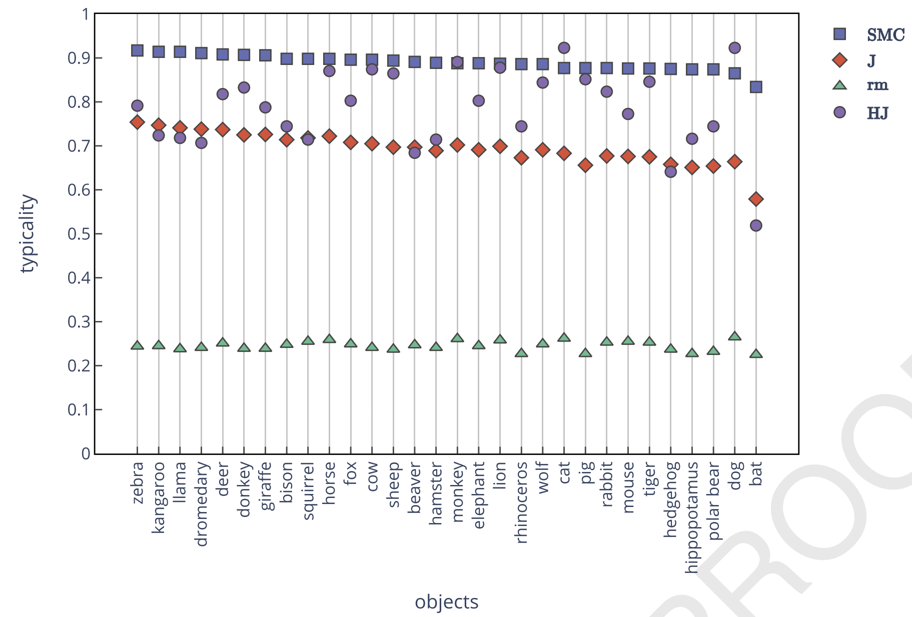

Fig. 3. Typicality ratings for "mammal" with objects ordered by values of typ_{SMC} and the values of typ_{HJ} rescaled to $[0, 1]$; AnimalCategory data.

association between two quantities, i.e. between two typicalities in our case. In particular, its value measures to what extent the ordering of exemplars in one list agrees with the ordering of exemplars in the other list. The coefficient ranges from 1 (same ordering) to -1 (inverse, i.e. opposite ordering). We used τ_b to account for ties in typicality values and used its implementation in a Python library [26].

Secondly, we used the $\tilde{\gamma}$ rank correlation coefficient of [8] (called the robust rank correlation coefficient by the authors). Namely, the Kendall tau only takes into account the orderings of exemplars and disregards the typicality degrees on which the ordering is based. One may object to this as follows. Consider three lists, each consisting of two objects x_1 and x_2, along with their typicalities, say

$$l_1 = \langle\langle x_1, 0.85\rangle, \langle x_2, 0.1\rangle\rangle,$$
$$l_2 = \langle\langle x_1, 0.85\rangle, \langle x_2, 0.8\rangle\rangle, \text{ and}$$
$$l_3 = \langle\langle x_2, 0.9\rangle, \langle x_1, 0.8\rangle\rangle.$$

The Kendall τ_b of l_1 and l_3 is -1 (opposite ordering), which is the same as τ_b of l_2 and l_3 (opposite ordering as well), since only the orderings matter. However, since we naturally also take the typicality degrees into account, l_2 and l_3 are much better correlated (since the typicality degrees are very close) than l_1 and l_3. The $\tilde{\gamma}$ coefficient resolves this by taking closeness of degrees into account. In particular, we set the parameter r, which controls what the method considers as close values of typicality, to $r = 0.2$.

Thirdly, we employed the idea put forward in our previous study [7] to alleviate the strictness of rank correlation consisting in basically looking solely at agreement of two compared orderings of objects. One might argue that rather than to examine agreement in ordering, it is more interesting to examine whether the set of the top r objects (i.e., r most typical objects) in one list is similar to the set of top r objects in the other list for various values of r. For a given typicality assignment M (e.g. M = typ_{J}), we denote the set of the top r objects in the list corresponding to M as

$$Top_r^{\mathrm{M}}.$$

For this, we assume that (a) if the $(r+1)$-st, …, $(r+k)$-th objects are tied with the r-th one, i.e. have the same value of typicality, we add these k objects to Top_r^{M}; (b) we do not include objects with typicality equal to 0. Now, given typicality assignments M and N, we are interested in whether and to what extent are the sets Top_r^{M} and Top_r^{N} similar. For this purpose, we proceed as follows. For objects x_1 and x_2, we denote by $s(x_1, x_2)$ a suitable defined similarity degree (a number in $[0, 1]$ in our case).

Below, we use

$$s(x_1, x_2) = sim_{\mathrm{J}}(\{x_1\}^{\uparrow}, \{x_2\}^{\uparrow}),$$

i.e. $s(x_1, x_2)$ equals the Jaccard index of the sets $\{x_1\}^{\uparrow}$ and $\{x_2\}^{\uparrow}$ of attributes for objects x_1 and x_2, respectively; cf. (4).

Table 4
Correlations of typicalities; AnimalCategory and AnimalExemplar data.

dataset	concept	τ_b correlation				$\tilde{\gamma}$ correlation			
AnimalCategory	bird		typ_{J}	typ_{rm}	typ_{HJ}		typ_{J}	typ_{rm}	typ_{HJ}
		typ_{SMC}	0.862	0.196	0.445	typ_{SMC}	0.978	0.426	0.634
		typ_{J}		0.334	0.5	typ_{J}		0.601	0.691
		typ_{rm}			0.319	typ_{rm}			0.496
	fish		typ_{J}	typ_{rm}	typ_{HJ}		typ_{J}	typ_{rm}	typ_{HJ}
		typ_{SMC}	0.919	0.531	0.412	typ_{SMC}	0.999	0.932	0.776
		typ_{J}		0.581	0.462	typ_{J}		0.93	0.758
		typ_{rm}			0.47	typ_{rm}			0.784
	mammal		typ_{J}	typ_{rm}	typ_{HJ}		typ_{J}	typ_{rm}	typ_{HJ}
		typ_{SMC}	0.871	0.014	-0.053	typ_{SMC}	0.981	0.042	-0.005
		typ_{J}		0.133	0.002	typ_{J}		0.28	0.115
		typ_{rm}			0.413	typ_{rm}			0.547
AnimalExemplar	bird		typ_{J}	typ_{rm}	typ_{HJ}		typ_{J}	typ_{rm}	typ_{HJ}
		typ_{SMC}	0.839	0.269	0.454	typ_{SMC}	0.968	0.398	0.652
		typ_{J}		0.43	0.56	typ_{J}		0.61	0.784
		typ_{rm}			0.505	typ_{rm}			0.717
	fish		typ_{J}	typ_{rm}	typ_{HJ}		typ_{J}	typ_{rm}	typ_{HJ}
		typ_{SMC}	0.927	0.428	0.333	typ_{SMC}	0.995	0.736	0.615
		typ_{J}		0.47	0.32	typ_{J}		0.749	0.553
		typ_{rm}			0.249	typ_{rm}			0.259
	mammal		typ_{J}	typ_{rm}	typ_{HJ}		typ_{J}	typ_{rm}	typ_{HJ}
		typ_{SMC}	0.582	-0.241	-0.345	typ_{SMC}	0.856	-0.449	-0.472
		typ_{J}		0.177	-0.002	typ_{J}		0.274	-0.025
		typ_{rm}			0.595	typ_{rm}			0.729

Finally, for two typicality assignments, M and N, and a given $r = 1, 2, 3, \ldots$, we define

$$S(Top_r^{\mathrm{M}}, Top_r^{\mathrm{N}}) = \min(I_{\mathrm{MN}}, I_{\mathrm{NM}})$$

where

$$I_{\mathrm{MN}} = \frac{\sum_{x_1 \in Top_r^{\mathrm{M}}} \max_{x_2 \in Top_r^{\mathrm{N}}} s(x_1, x_2)}{|Top_r^{\mathrm{M}}|}$$

and, symmetrically,

$$I_{\mathrm{NM}} = \frac{\sum_{x_2 \in Top_r^{\mathrm{N}}} \max_{x_1 \in Top_r^{\mathrm{M}}} s(x_1, x_2)}{|Top_r^{\mathrm{N}}|}.$$

According to basic rules of fuzzy logic, $S(Top_r^{\mathrm{M}}, Top_r^{\mathrm{N}})$ may naturally be interpreted as the truth degree of the proposition "for most objects in Top_r^{M} there is a similar object in Top_r^{N} and vice versa." Due to this interpretation and since S is actually a reflexive and symmetric fuzzy relation [3,28], S is a good candidate for measuring similarity of sets of objects. High values of S indicate high similarity and $S(Top_r^{\mathrm{M}}, Top_r^{\mathrm{N}}) = 1$ takes place if and only if $Top_r^{\mathrm{M}} = Top_r^{\mathrm{N}}$.

Results of analyses

Consider first the rank correlations τ_b and $\tilde{\gamma}$, which are shown for the concepts "bird," "fish," and "mammal" in Table 4. The table presents, both for the AnimalCategory and AnimalExemplar data, and for the three concepts in this data all the six correlation coefficients τ_b and six correlation coefficients $\tilde{\gamma}$ for the four observed typicalities typ_{SMC}, typ_{J}, typ_{rm}, and typ_{HJ}.

Let us first examine the τ_b correlations of the three computed typicalities with human judgment. Note first that according to a commonly accepted interpretation, the values of τ_b may be interpreted as follows: $\tau_b \geq 0.3$, $0.2 \leq \tau_b < 0.3$, $0.1 \leq \tau_b < 0.2$, and $0.0 \leq \tau_b < 0.1$ indicate strong, moderate, weak and very weak correlation, respectively (analogously for negative values). All the computed typicalities, typ_{SMC}, typ_{J}, typ_{rm} display a strong correlation with human judgment in virtually all cases except for the concept "mammal". For this concept, only typ_{rm} exhibits a strong correlation. We only have a partial explanation for this. Namely, we believe that "mammal" is a somewhat problematic concept as regards human judgment of typicality (we observed this when collecting human rankings for "mammal" in the Zoo data; see the next section); we encountered similar difficulties with some other concepts, e.g. "kitchen utensils" (what is a typical kitchen utensil?). Why some concepts are problematic in this sense is a question that should be explored in the future, possibly with the help of psychologists.

Next, let us examine mutual correlations of the three computed typicalities. The data indicates a very strong correlation between typ_{SMC} and typ_{J} in most cases. Furthermore, we see significantly less strong correlations between typ_{SMC} and typ_{rm}, between typ_{J} and typ_{rm}; yet these correlations range from moderate to strong in most cases, except for the above-discussed "mammal".

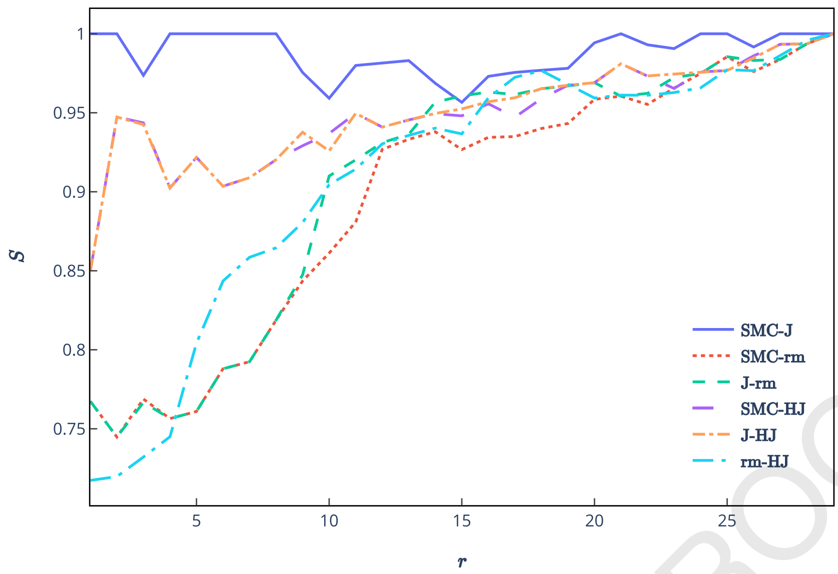

Fig. 4. Similarity S of top r typical objects; concept “bird” in AnimalCategory data. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

Similar pattern may be observed for the $\tilde{\gamma}$ correlations. Since we observed by and large the same behavior for all the data and concepts we examined, including the Zoo data presented in the next section, the correlation analysis suggests as interesting the problem to analyze, both experimentally and theoretically, the relationships of the three computed typicalities and, in particular, to focus on why typ_{rm} seems to perform differently from the two rather correlated ones.

Let us next consider our third method of comparison. The mutual similarities of the sets of top r typical objects for the AnimalCategory and the AnimalExemplar data for our three examined concepts are shown in Figs. 4–9. Each figure displays six graphs representing the six mutual similarities $S = S(Top_r^{\mathrm{M}}, Top_r^{\mathrm{N}})$ of the sets of top r typical objects according to typicality M and typicality N (vertical axis), for increasing $r = 1, 2, \ldots$ (horizontal axis). The graphs reveal a similar pattern of relationships we observed with the correlations. For instance, in Fig. 4, the blue graph labeled SMC-J representing the similarities for M = $\mathrm{typ}_{\mathrm{SMC}}$ and N = $\mathrm{typ}_{\mathrm{J}}$ shows a high similarity of the sets of top r $\mathrm{typ}_{\mathrm{SMC}}$-typical and top r $\mathrm{typ}_{\mathrm{J}}$-typical birds, which confirms—from a different perspective—the very strong rank correlations of these two typicalities observed above. The two lines, one for M = $\mathrm{typ}_{\mathrm{SMC}}$ and N = $\mathrm{typ}_{\mathrm{HJ}}$, the other for M = $\mathrm{typ}_{\mathrm{J}}$ and N = $\mathrm{typ}_{\mathrm{HJ}}$, which also attain high values even for small r confirm strong correlations of these two pairs of typicalities observed above. Naturally, the similarities increase with increasing r which needs to be taken into account when interpreting the graphs.

4.2. Experiments with zoo data

Zoo data

Zoo is a commonly known dataset [12] and its concepts are mostly well interpretable. It describes 101 animals (objects) by their 17 attributes and has the density of 0.36. We remove the somewhat disputable object “girl” from the data; we renamed one of the two objects denoted “frog” to “frog venomous.” All of the attributes are yes/no attributes except for the attribute describing the number of legs, which we nominally scaled, and an attribute determining the type of animal, which we removed. The scaled data is presented in Table 5 (to save space, objects with the same attributes are put on the same row).

The Zoo data (i.e. the formal context corresponding to the data) contains several formal concepts, among them three concepts that may be interpreted as “bird,” “fish,” and “mammal”. Since concepts with the same interpretation (in different data, however, and thus represented by different sets of objects and attributes) were used in the previous section, we examine typicalities for these concepts.

To be able to perform similar analyses to those we described the previous section, we first obtained human typicality ratings for the objects of the three concepts by means of a questionnaire; see [4] for the data and appendix for the questionnaire. Since the obtained data may be useful for further studies, we describe it to a certain detail. Altogether, 242 respondents participated in the survey. We first split the respondents in four groups (students at Palacky University Olomouc, our coworkers, relatives, and others), since we assumed possibly different reliability of these groups. However, as a correlation analysis revealed high correlations between the average ratings in these groups (Kendall τ_b always higher than 0.6), we merged the groups into a single group for which we computed average typicality ratings. Note also that approximately 56% (136) of women and 44% (106) of men were among the participants. Median, minimum and maximum age of participants was 23, 17 and 81. The three concepts along with the obtained human judgment of typicality and the three

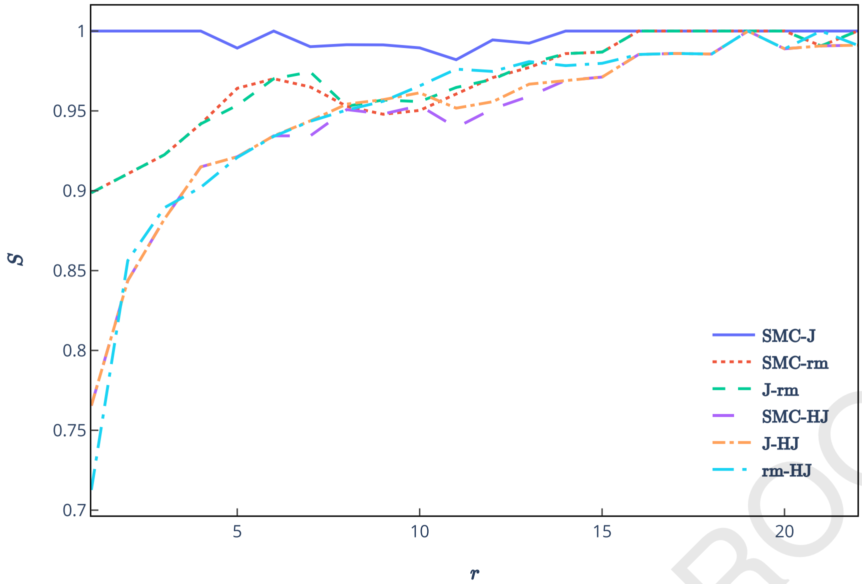

Fig. 5. Similarity S of top r typical objects; concept “fish” in AnimalCategory data.

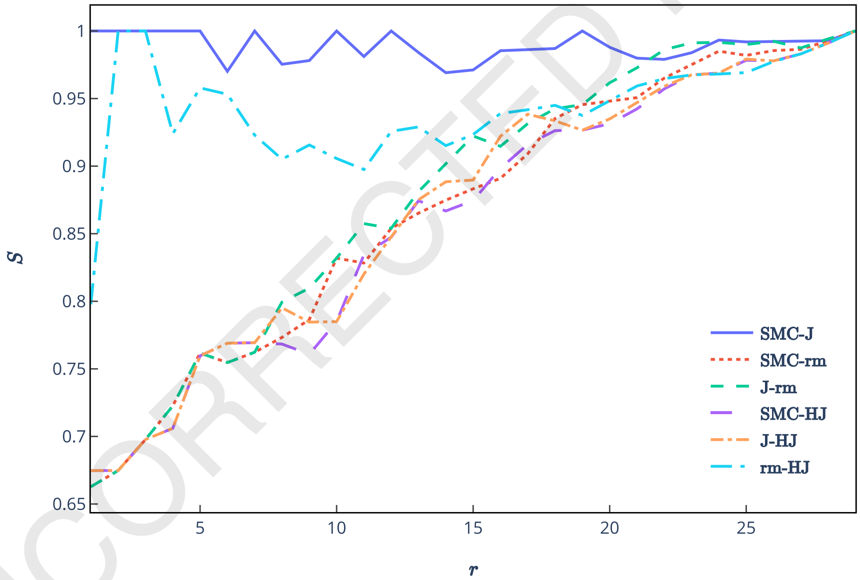

Fig. 6. Similarity S of top r typical objects; concept “mammal” in AnimalCategory data.

computed typicality ratings are shown in Tables 6, 7, and 8, respectively, in the same manner as with the Dutch data. As in the previous section, the typicality data is also displayed in Figs. 10, 11, and 12.

Results of analyses

The correlations of the typicalities for the three concepts are shown in Table 9. The mutual similarities of the set of top r objects by the observed typicalities are displayed in Figs. 13, 14, and 15.

Basically, a similar pattern as for the Dutch data may be observed with the following differences. First, correlations of the computed typicalities to human judgment are generally somewhat smaller, which is due to the fact that the attributes in the Zoo data are considerably less informative compared to the attributes in Dutch data (several animals have the same attributes in the Zoo data but, at the same time, are rather different as regards typicality). This illustrates the need of informative attributes for the computed typicalities to work reasonably well, as mentioned above. In addition, typ_{rm}, which again behaves somewhat differently compared to the strongly correlated typ_{SMC} and typ_{J}, achieves smaller correlation with hu-

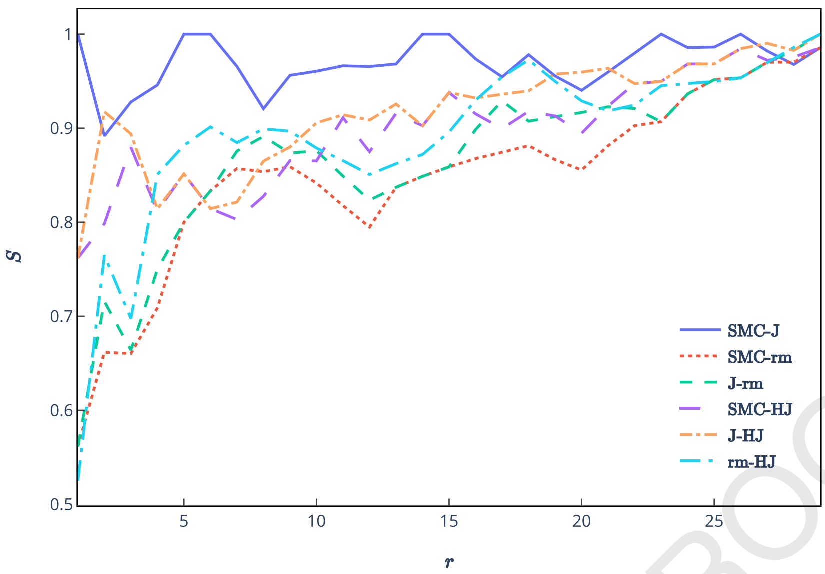

Fig. 7. Similarity S of top r typical objects; concept "bird" in AnimalExemplar data.

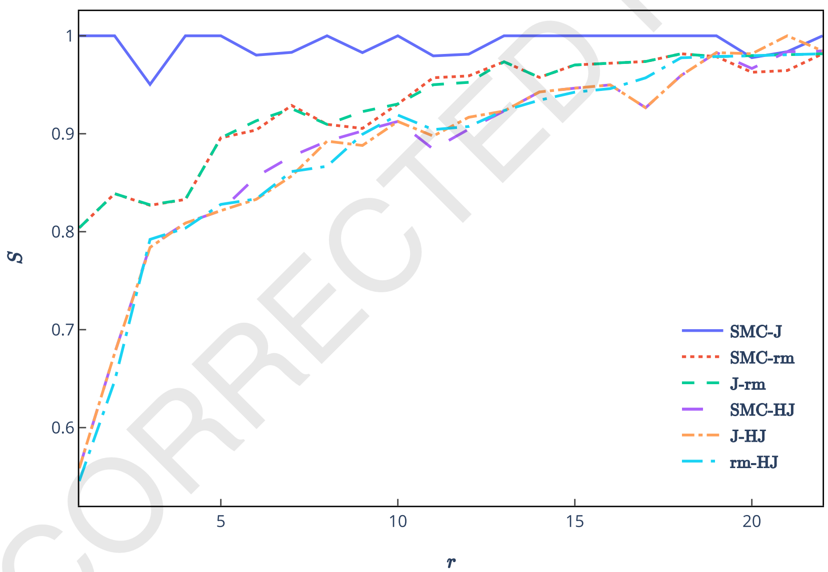

Fig. 8. Similarity S of top r typical objects; concept "fish" in AnimalExemplar data.

man judgment compared to Dutch data. This again calls for a closer examination of the relationship between the computed typicalities.

Finally, let us mention an instructive observation regarding the concept "fish." Here, carp appears to be a problematic exemplar. While most of our respondents consider it the most typical fish, it is considered atypical according to our formulas. The reason is, on the one hand, that the presence of the attribute domestic and absence of predator make carp atypical according to the typicality formulas. On the other hand, the perception of carp by respondents is influenced by other factors, including cultural background, which—as in this case—may be considerably more significant for human judgment than the actual attributes present in the data when it comes to determination of typicality. This phenomenon is discussed in the literature [20] and calls for a closer examination from a general perspective. Note also that the presence of this single problematic exemplar makes the correlations of the computed typicalities with human judgment very weak. When carp is removed from the extent of the concept "fish" (in which case the set of objects no longer forms an extent of a formal concept), the rank correlations with human judgment are very strong; see the correlations in the part "fish (without carp)".

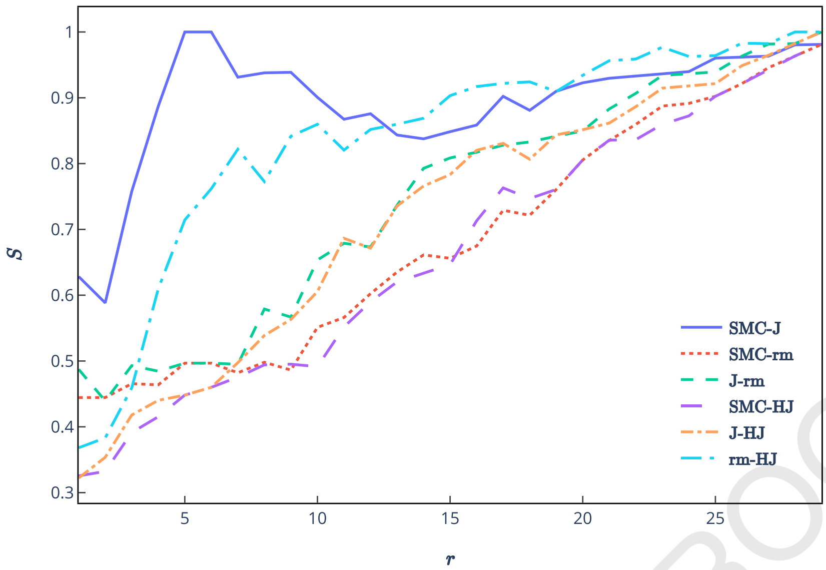

Fig. 9. Similarity S of top r typical objects; concept "mammal" in AnimalExemplar data.

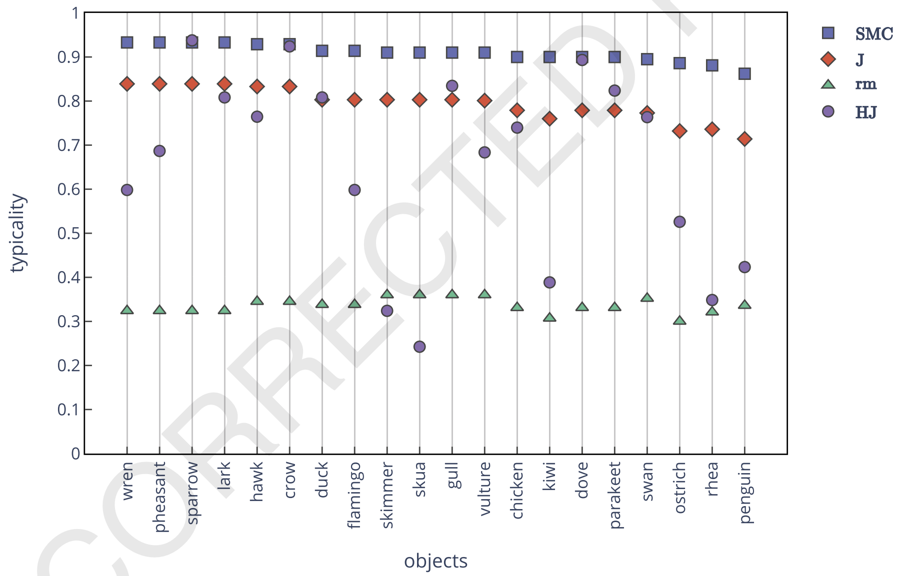

Fig. 10. Typicality ratings for "bird" with objects ordered by values of typ_{SMC} and the values of typ_{HJ} rescaled to $[0, 1]$; Zoo data.

5. Conclusions and further topics

This paper is intended to make first steps in studying and exploiting typicality within a more formalized setting compared to what is common in the literature on the psychology of concepts. The basic aim is to enable a more precise analysis of typicality, both experimental and theoretical.

We proposed a formal definition of typicality, which translates to a general scheme for a formula to compute typicality of objects of a given concept in a given data consisting of objects, attributes and an incidence relation between objects and attributes. Our scheme is based on a basic psychological view of typicality due to Rosch and Mervis. We considered three typicality functions resulting from the general scheme, namely typ_{SMC}, typ_{J}, and typ_{rm}, the last of which was proved equivalent to a function actually proposed by Rosch and Mervis. Experiments performed with the Dutch and the Zoo data revealed that for most concepts in this data for which human judgment on typicality is available, there is a strong agreement between the computed typicalities and human judgment. This finding was confirmed by three kinds of analyses. We also observed a very strong correlation of the functions typ_{SMC} and typ_{J}, and a considerably weaker correlation of these functions

Table 5
Scaled Zoo data.

	hair	feathers	eggs	milk	airborne	aquatic	predator	toothed	backbone	breathes	venomous	fins	tail	domestic	catsize	legs 0	legs 2	legs 4	legs 5	legs 6	legs 8
scorpion							×			×	×		×								×
seasnake						×	×	×	×		×		×			×					
dolphin, porpoise				×		×	×	×	×	×		×	×		×	×					
flea, termite			×							×										×	
slug, worm			×							×						×					
tortoise			×						×	×			×		×			×			
clam			×				×									×					
tuatara			×				×	×	×	×			×					×			
slowworm			×				×	×	×	×			×			×					
pitviper			×				×	×	×	×	×		×			×					
haddock, seahorse, sole			×			×		×	×			×	×			×					
carp			×			×		×	×			×	×	×		×					
toad			×			×		×	×	×								×			
crayfish, lobster			×			×	×													×	
starfish			×			×	×												×		
crab			×			×	×											×			
octopus			×			×	×								×						×
seawasp			×			×	×				×					×					
bass, catfish, chub, herring, piranha			×			×	×	×	×			×	×			×					
dogfish, pike, tuna			×			×	×	×	×			×	×		×	×					
stingray			×			×	×	×	×		×	×	×		×	×					
frog			×			×	×	×	×	×								×			
newt			×			×	×	×	×	×			×					×			
frog venomous			×			×	×	×	×	×	×							×			
gnat			×		×					×										×	
ladybird			×		×		×			×										×	
ostrich		×	×						×	×			×		×		×				
kiwi		×	×				×		×	×			×				×				
rhea		×	×				×		×	×			×		×		×				
penguin		×	×			×	×		×	×			×		×		×				
lark, pheasant, sparrow, wren		×	×		×				×	×			×				×				
flamingo		×	×		×				×	×			×		×		×				
chicken, dove, parakeet		×	×		×				×	×			×	×			×				
crow, hawk		×	×		×		×		×	×			×				×				
vulture		×	×		×		×		×	×			×		×		×				
duck		×	×		×	×			×	×			×				×				
swan		×	×		×	×			×	×			×		×		×				
gull, skimmer, skua		×	×		×	×	×		×	×			×				×				
gorilla	×			×				×	×	×					×		×				
cavy	×			×				×	×	×				×				×			
hare, vole	×			×				×	×	×			×					×			
squirrel	×			×				×	×	×			×				×				
antelope, buffalo, deer, elephant, giraffe, oryx	×			×				×	×	×			×		×			×			
wallaby	×			×				×	×	×			×		×		×				
hamster	×			×				×	×	×			×	×				×			
calf, goat, pony, reindeer	×			×				×	×	×			×	×	×			×			
aardvark, bear	×			×			×	×	×	×					×			×			
mole, opossum	×			×			×	×	×	×			×					×			
pussycat	×			×			×	×	×	×			×	×	×			×			
mink	×			×		×	×	×	×	×			×		×			×			
seal	×			×		×	×	×	×	×		×			×	×					
sealion	×			×		×	×	×	×	×		×	×		×		×				
fruitbat, vampire	×			×	×			×	×	×			×				×				
housefly, moth	×		×		×					×										×	
wasp	×		×		×					×	×									×	
honeybee	×		×		×					×	×			×						×	
platypus	×		×	×		×	×		×	×			×		×			×			
boar, cheetah, leopard, lion, lynx, mongoose	×			×			×	×	×	×			×		×			×			
polecat, puma, raccoon, wolf	×			×			×	×	×	×			×		×			×			

with typ_{rm}, which implies a need for a further detailed analysis of the proposed typicality functions. We also pointed out problems to consider from a psychological viewpoint which are relevant for evaluation of formal definitions of typicality.

Major topics to be explored in the future are the following:

Table 6
Typicality ratings for "bird"; Zoo data.

typ_{SMC} order	typ_{SMC}	typ_{J} order	typ_{J}	typ_{rm} order	typ_{rm}	typ_{HJ} order	typ_{HJ}
wren	0.933	wren	0.839	vulture	0.360	sparrow	4.751
pheasant	0.933	pheasant	0.839	gull	0.360	crow	4.696
sparrow	0.933	sparrow	0.839	skua	0.360	dove	4.573
lark	0.933	lark	0.839	skimmer	0.360	gull	4.338
hawk	0.929	hawk	0.833	swan	0.352	parakeet	4.295
crow	0.929	crow	0.833	hawk	0.345	duck	4.233
duck	0.914	duck	0.803	crow	0.345	lark	4.233
flamingo	0.914	flamingo	0.803	duck	0.338	hawk	4.058
vulture	0.910	gull	0.803	flamingo	0.338	swan	4.054
gull	0.910	skua	0.803	penguin	0.336	chicken	3.959
skua	0.910	skimmer	0.803	chicken	0.331	pheasant	3.747
skimmer	0.910	vulture	0.801	parakeet	0.331	vulture	3.734
chicken	0.900	chicken	0.779	dove	0.331	flamingo	3.393
kiwi	0.900	dove	0.779	pheasant	0.324	wren	3.393
dove	0.900	parakeet	0.779	lark	0.324	ostrich	3.104
parakeet	0.900	swan	0.773	sparrow	0.324	penguin	2.693
swan	0.895	kiwi	0.760	wren	0.324	kiwi	2.554
ostrich	0.886	rhea	0.736	rhea	0.321	rhea	2.394
rhea	0.881	ostrich	0.732	kiwi	0.307	skimmer	2.296
penguin	0.862	penguin	0.714	ostrich	0.300	skua	1.970

Table 7
Typicality ratings for "fish"; Zoo data.

typ_{SMC} order	typ_{SMC}	typ_{J} order	typ_{J}	typ_{rm} order	typ_{rm}	typ_{HJ} order	typ_{HJ}
bass	0.963	bass	0.913	stingray	0.385	carp	4.860
catfish	0.963	catfish	0.913	dogfish	0.381	pike	4.711
chub	0.963	chub	0.913	pike	0.381	catfish	4.513
herring	0.963	herring	0.913	tuna	0.381	haddock	4.430
piranha	0.963	piranha	0.913	bass	0.366	bass	4.429
dogfish	0.945	dogfish	0.875	catfish	0.366	tuna	4.418
haddock	0.945	pike	0.875	chub	0.366	piranha	4.054
pike	0.945	tuna	0.875	herring	0.366	herring	4.018
seahorse	0.945	haddock	0.868	piranha	0.366	dogfish	3.104
sole	0.945	seahorse	0.868	carp	0.337	chub	2.992
tuna	0.945	sole	0.868	haddock	0.333	stingray	2.948
carp	0.905	stingray	0.803	seahorse	0.333	seahorse	2.242
stingray	0.905	carp	0.788	sole	0.333	sole	2.042

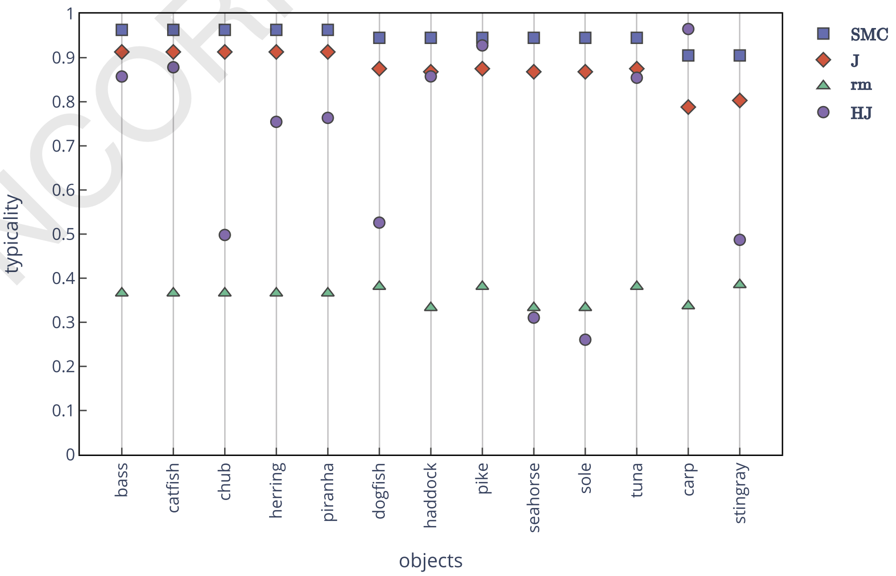

Fig. 11. Typicality ratings for "fish" with objects ordered by values of typ_{SMC} and the values of typ_{HJ} rescaled to [0, 1]; Zoo data.

Table 8
Typicality ratings for “mammal”; Zoo data.

typ_{SMC} order	typ_{SMC}	typ_{J} order	typ_{J}	typ_{rm} order	typ_{rm}	typ_{HJ} order	typ_{HJ}
wolf	0.912	wolf	0.822	pussycat	0.383	pussycat	4.714
cheetah	0.912	cheetah	0.822	mink	0.382	calf	4.636
lynx	0.912	lynx	0.822	leopard	0.375	goat	4.521
lion	0.912	mongoose	0.822	raccoon	0.375	bear	4.438
leopard	0.912	polecat	0.822	puma	0.375	gorilla	4.423
polecat	0.912	leopard	0.822	polecat	0.375	wolf	4.397
puma	0.912	lion	0.822	mongoose	0.375	hare	4.328
raccoon	0.912	puma	0.822	lynx	0.375	lion	4.278
boar	0.912	boar	0.822	lion	0.375	deer	4.269
mongoose	0.912	raccoon	0.822	wolf	0.375	pony	4.088
deer	0.910	buffalo	0.810	boar	0.375	boar	4.059
elephant	0.910	elephant	0.810	cheetah	0.375	cheetah	4.033
giraffe	0.910	oryx	0.810	calf	0.358	elephant	4.021
oryx	0.910	giraffe	0.810	goat	0.358	lynx	3.967
buffalo	0.910	antelope	0.810	reindeer	0.358	giraffe	3.942
antelope	0.910	deer	0.810	pony	0.358	leopard	3.934
opossum	0.886	pussycat	0.774	sealion	0.357	squirrel	3.908
mole	0.886	mink	0.765	elephant	0.350	puma	3.900
hare	0.883	pony	0.762	giraffe	0.350	hamster	3.892
vole	0.883	goat	0.762	buffalo	0.350	cavy	3.863
pussycat	0.881	reindeer	0.762	deer	0.350	reindeer	3.801
pony	0.879	calf	0.762	antelope	0.350	antelope	3.791
reindeer	0.879	mole	0.760	oryx	0.350	buffalo	3.770
mink	0.879	opossum	0.760	mole	0.338	wallaby	3.628
goat	0.879	hare	0.746	opossum	0.338	raccoon	3.619
calf	0.879	vole	0.746	platypus	0.337	vole	3.608
bear	0.876	bear	0.741	bear	0.333	polecat	3.404
aardvark	0.876	aardvark	0.741	aardvark	0.333	mink	3.311
hamster	0.852	hamster	0.703	hamster	0.321	mole	3.302
wallaby	0.850	wallaby	0.690	wallaby	0.320	dolphin	3.186
squirrel	0.824	squirrel	0.631	hare	0.313	opossum	3.172
cavy	0.817	sealion	0.622	vole	0.313	oryx	3.016
gorilla	0.814	cavy	0.621	seal	0.312	seal	2.917
platypus	0.788	platypus	0.620	porpoise	0.308	aardvark	2.899
fruitbat	0.781	gorilla	0.610	dolphin	0.308	sealion	2.837
sealion	0.781	fruitbat	0.581	fruitbat	0.286	fruitbat	2.474
vampire	0.781	vampire	0.581	vampire	0.286	mongoose	2.438
seal	0.738	seal	0.549	squirrel	0.283	porpoise	2.425
porpoise	0.731	porpoise	0.540	cavy	0.280	vampire	2.242
dolphin	0.731	dolphin	0.540	gorilla	0.279	platypus	2.230

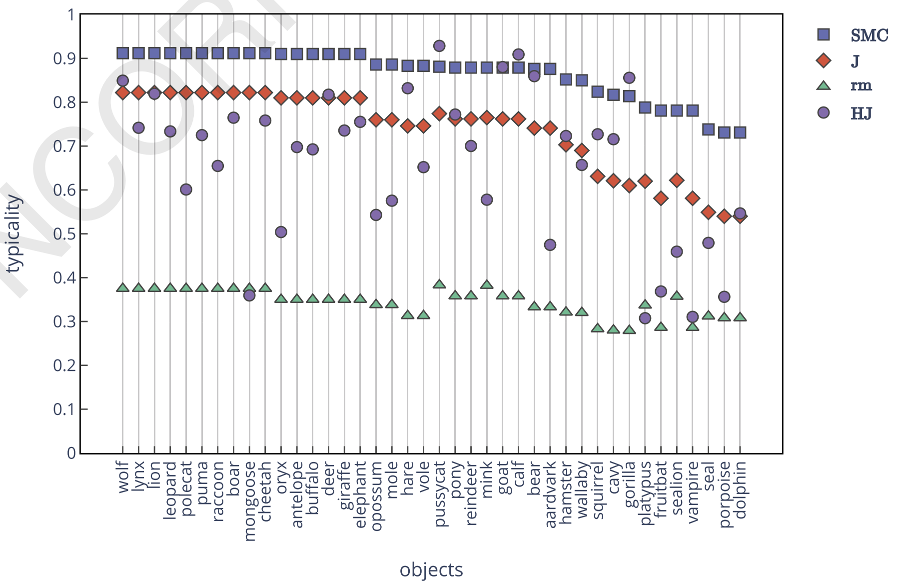

Fig. 12. Typicality ratings for “mammal” with objects ordered by values of typ_{SMC} and the values of typ_{HJ} rescaled to $[0, 1]$; Zoo data.

Table 9
Correlations of typicalities; Zoo data.

dataset	concept	τ_b correlation				$\tilde{\gamma}$ correlation			
Zoo	bird		typ_J	typ_{rm}	typ_{HJ}		typ_J	typ_{rm}	typ_{HJ}
		typ_{SMC}	0.948	0.012	0.267	typ_{SMC}	0.995	0.073	0.45
		typ_J		0.063	0.264	typ_J		0.229	0.468
		typ_{rm}			0.044	typ_{rm}			0.057
	fish		typ_J	typ_{rm}	typ_{HJ}		typ_J	typ_{rm}	typ_{HJ}
		typ_{SMC}	0.916	-0.106	0.0	typ_{SMC}	1.0	0.088	-0.059
		typ_J		0.065	0.058	typ_J		0.273	-0.074
		typ_{rm}			0.029	typ_{rm}			0.132
	fish (without carp)		typ_J	typ_{rm}	typ_{HJ}		typ_J	typ_{rm}	typ_{HJ}
		typ_{SMC}	0.995	0.831	0.556	typ_{SMC}	0.995	0.831	0.556
		typ_J		0.873	0.509	typ_J		0.873	0.509
		typ_{rm}			0.296	typ_{rm}			0.296
	mammal		typ_J	typ_{rm}	typ_{HJ}		typ_J	typ_{rm}	typ_{HJ}
		typ_{SMC}	0.904	0.599	0.238	typ_{SMC}	0.995	0.831	0.556
		typ_J		0.695	0.27	typ_J		0.873	0.482
		typ_{rm}			0.233	typ_{rm}			0.296

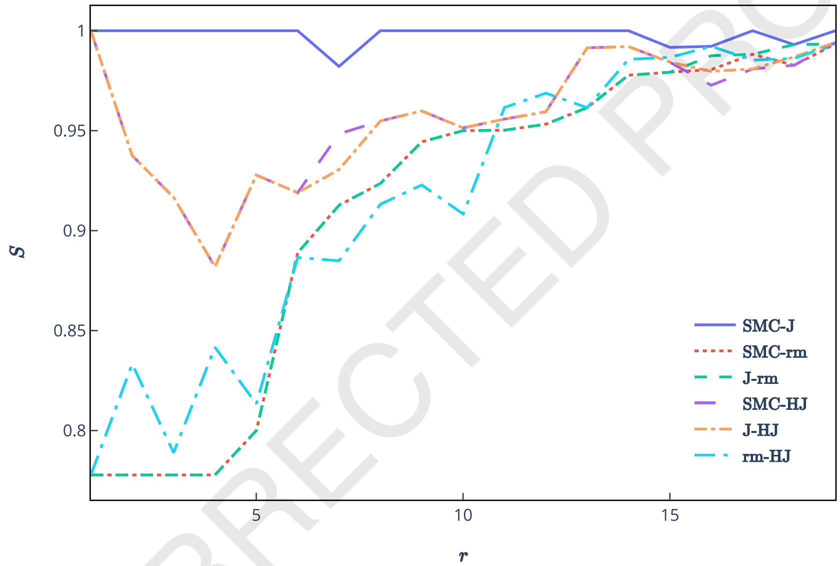

Fig. 13. Similarity S of top r typical objects; concept "bird" in Zoo data.

- Crucial for performing experimental evaluation of formal approaches to typicality and surrounding phenomena is availability of quality data that includes data describing human judgment. Dutch data seems to be the most comprehensive available data for this purpose. Obtaining such data and making the data publicly available appears to be an important goal.
- Our experience with obtaining human judgment of typicality suggests that data describing human judgment should not be taken as "ground truth" (to use a term often mentioned in the psychological literature), for which a perfect fit is required with a given formula for computing typicality. Namely, the data on human judgment may have its own issues some of which are mentioned above. Rather than seeking a perfect fit, an evaluation of a proposed formalization of typicality needs to be performed with caution. This issue seems to point out an important methodological question which involves both psychological and mathematical aspects.
- As mentioned above, our experiments suggest that it is important to analyze, experimentally and theoretically, further relationships between the three proposed typicality functions, as well as to explore further instances of our general scheme for typicality formulas. In particular, explorations of further similarity functions is needed. As an example, we performed experiments with similarity functions which disregard attributes from the intent of the concept for which typicality is evaluated; this approach seems to have some advantages over the three similarity functions described above, such as better distinction between typical and atypical objects.
- Formalization of other existing psychological views of typicality, such as those mentioned in section 2.2, clearly represents a related, important goal. This includes the possibility to take into account the second part of Rosch and Mervis view of typicality, mentioned in section 2.2, which regards similarity to objects in other categories. More radical depar-

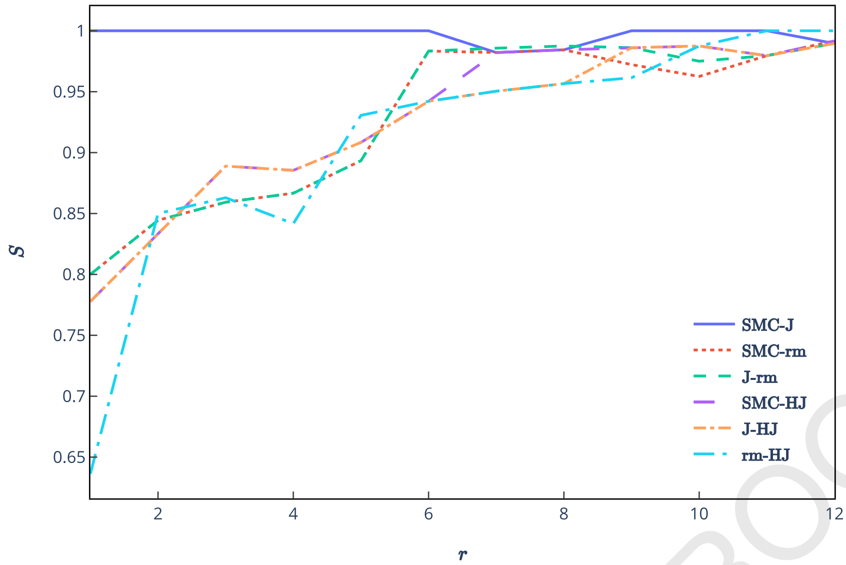

Fig. 14. Similarity S of top r typical objects; concept "fish" in Zoo data.

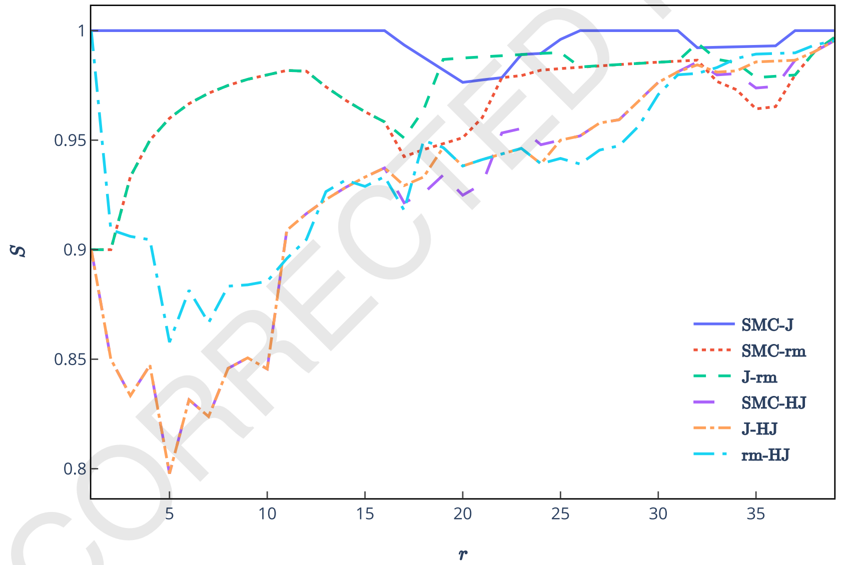

Fig. 15. Similarity S of top r typical objects; concept "mammal" in Zoo data.

tures from our present approach would take dependence of typicality on context into account (in a general sense of the notion of context), as suggested, e.g. in [27] and [10].

- In addition to typicality of objects, typicality of attributes may be explored. At the first sight, this seems just a dual case of typicality of objects. From a psychological point of view, however, typicality of attributes has a rather different role; see e.g. [20]. Methods to determine typicality of attributes shall thus be explored.
- Due to considerable cognitive significance of typicality, it seems natural to exploit typicality in machine learning and data analysis and thus extend the existing attempts mentioned in section 3.2.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This paper is an extended version of Belohlavek, R., and Mikula, T., "Typicality in conceptual structures within the framework of formal concept analysis," Proc. CLA 2020, 33–45. We thank anonymous reviewers for suggesting useful improvements in the conference version of this paper. Supported partly by the project IGA 2020, reg. no. IGA_PrF_2020_019, and by the project IGA 2021, reg. no. IGA_PrF_2021_022, of Palacký University, Olomouc.

Appendix

Instructions of the questionnaire we used to collect the typicality ratings for the Zoo data (the original questionnaire was in Czech):

Hello!
By filling out this questionnaire you contribute to research in the psychology of concepts at the Department of Computer Science, Palacky University Olomouc. The questionnaire takes cca 10 minutes.
You will be asked to assess typicality of animals for three categories (concepts), namely bird, mammal, and fish. Each category shall be assessed on a separate page.
For a given category (e.g. bird), you will see a list of particular animals (birds) of this category. Read the whole list first. Then select for each animal a value in the scale 1 to 5 which describes the extent to which the animal is typical of the category (1 = least typical, 5 = most typical). If you consider it necessary, when filling out the values, go back and change the previously filled values. If you do not know the particular animal, do not select any value (go to the next animal).
When filling out the questionnaire, do not search for additional information (e.g. pictures). Do not spend much time when assigning a typicality value (cca seconds). Do not forget to send out the filled questionnaire. Fill the questionnaire just once.
After sending out your questionnaire, you will be able to see responses of other respondents.
Do not hesitate to contact us if you have questions. Thank you for filling out the questionnaire.

References

[1] L.W. Barsalou, Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories, J. Exp. Psychol. Learn. Mem. Cogn. 11 (4) (1985) 629–654, https://doi.org/10.1037/0278-7393.11.1-4.629.
[2] R. Belohlavek, Fuzzy Relational Systems: Foundations and Principles, Kluwer Academic Publishers, USA, 2002.
[3] R. Belohlavek, J.W. Dauben, G.J. Klir, Fuzzy Logic and Mathematics: A Historical Perspective, Oxford University Press, 2017.
[4] R. Belohlavek, T. Mikula, Zoo typicality, GitHub repository, 2021, https://github.com/mikulatomas/zoo-typicality.
[5] R. Belohlavek, M. Trnecka, Basic level in formal concept analysis: interesting concepts and psychological ramifications, in: Twenty-Third International Joint Conference on Artificial Intelligence, 2013, pp. 1233–1239.
[6] R. Belohlavek, M. Trnecka, Basic level of concepts in formal concept analysis 1: formalization and utilization, Int. J. Gen. Syst. 49 (7) (2020) 689–706.
[7] R. Belohlavek, M. Trnecka, Basic level of concepts and formal concept analysis 2: examination of existing basic level metrics, Int. J. Gen. Syst. 49 (7) (2020) 707–723.
[8] U. Bodenhofer, F. Klawonn, Robust rank correlation coefficients on the basis of fuzzy orderings: initial steps, Mathw. Soft Comput. 15 (2008) 5–20.
[9] C. Carpineto, G. Romano, Concept Data Analysis: Theory and Applications, Wiley, Chichester, England, Hoboken, NJ, 2004.
[10] M.A. Dieciuc, J.R. Folstein, Typicality: stable structures and flexible functions, Psychon. Bull. Rev. 26 (2019) 491–505, https://doi.org/10.3758/s13423-018-1546-2.
[11] S. De Deyne, S. Verheyen, E. Ameel, W. Vanpaemel, M.J. Dry, W. Voorspoels, G. Storms, Exemplar by feature applicability matrices and other Dutch normative data for semantic concepts, Behav. Res. Methods 40 (4) (2008) 1030–1048, https://doi.org/10.3758/BRM.40.4.1030.
[12] D. Dua, C. Graff, UCI Machine Learning Repository, University of California, Irvine, School of Information and Computer Sciences, 2019, http://archive.ics.uci.edu/ml.
[13] B.S. Everitt, S. Landau, M. Leese, D. Stahl, Cluster Analysis, 5th ed., Wiley, Chichester, West Sussex, UK, 2011.
[14] D.H. Fisher, A computational account of basic level and typicality effects, in: Proc. AAAI-88, 1988, pp. 233–238.
[15] B. Ganter, R. Wille, Formal Concept Analysis: Mathematical Foundations, Springer, Berlin, 1999.
[16] P. Jaccard, Nouvelles recherches sur la distribution florale, Bull. Soc. Vaud. Sci. Nat. 44 (1908) 223–370.
[17] M.-J. Lesot, R. Kruse, Data summarisation by typicality-based clustering for vectorial and non vectorial data, in: FUZZ-IEEE 2006, 2006, pp. 547–554.
[18] E. Machery, 100 years of psychology of concepts: the theoretical notion of concept and its operationalization, Stud. Hist. Philos. Biol. Biomed. Sci. 38 (1) (2007) 63–84, https://doi.org/10.1016/j.shpsc.2006.12.005.
[19] C.B. Mervis, J. Catlin, E. Rosch, Relationships among goodness-of-example, category norms, and word frequency, Bull. Psychon. Soc. 7 (3) (1976) 283–284, https://doi.org/10.3758/BF03337190.
[20] G.L. Murphy, The Big Book of Concepts, MIT Press, Cambridge, Mass, 2002.
[21] H.-S. Park, C.-H. Jun, A simple and fast algorithm for K-medoids clustering, Expert Syst. Appl. 36 (2) (2009) 3336–3341.
[22] E. Rosch, Principles of categorization, in: E. Rosch, B.B. Lloyd (Eds.), Cognition and Categorization, Erlbaum, Hillsdale, NJ, 1978, pp. 27–48.
[23] E. Rosch, C.B. Mervis, Family resemblances: studies in the internal structure of categories, Cogn. Psychol. 7 (4) (1975) 573–605, https://doi.org/10.1016/0010-0285(75)90024-9.
[24] E. Rosch, et al., Basic objects in natural categories, Cogn. Psychol. 8 (3) (1976) 382–439, https://doi.org/10.1016/0010-0285(76)90013-X.
[25] J.P. Rousseeuw, Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, Comput. Appl. Math. 20 (1987) 53–65, https://doi.org/10.1016/0377-0427(87)90125-7.
[26] SciPy 1.0 Contributors, et al., SciPy 1.0: fundamental algorithms for scientific computing in Python, Nat. Methods (2020), https://doi.org/10.1038/s41592-019-0686-2.
[27] W. Yeh, L.W. Barsalou, The situated nature of concepts, Am. J. Psychol. 119 (3) (2006) 349–384, https://doi.org/10.2307/20445349.
[28] L.A. Zadeh, Fuzzy sets, Inf. Control 8 (3) (1965) 338–353, https://doi.org/10.1016/S0019-9958(65)90241-X.
[29] J. Zhang, Selecting typical instances in instance-based learning, in: Machine Learning Proceedings, Morgan Kaufman, 1992, pp. 470–479.