Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Biomedical Informatics xxx (2005) xxx–xxx
Journal of Biomedical Informatics
www.elsevier.com/locate/yjbin

Methodological review

Conceptualizing the world: Lessons from history

Alexa T. McCray

Center for Clinical Computing, BIDMC, Harvard Medical School, Department of Medicine, 330 Brookline Avenue, Boston, MA 02215, USA

Received 21 June 2005

Abstract

Throughout history, philosophers, scientists, and other scholars have named and organized the salient elements of the world. These efforts have led to conceptualizations that differ widely in both content and form. This paper argues that all conceptualizations are biased, both because they depend on the purposes for which they have been created, and because they are closely tied to the world view of their designers. This bias needs to be recognized, and its consequences need to be addressed if the conceptualizations are to be used for purposes other than those for which they were designed. The paper begins with a brief overview of the disciplines that have been concerned with conceptualizing particular domains. This is followed by a non-exhaustive, illustrative, historical perspective, and the paper concludes with the interesting case study of biological taxonomy.

Keywords: Metaphysics; Classification; Knowledge representation; Semantics; Terminology; Concept formation; History

1. Introduction

For over two millennia there have been efforts by philosophers, scientists, and other scholars to name and organize the salient elements of the world as they understood it. These efforts have led to conceptualizations that differ not only in their content and form, but also, and, importantly, in the purpose for which they were developed. Gruber has defined a conceptualization as "...the objects, concepts, and other entities that are assumed to exist in some area of interest and the relationships that hold among them... A conceptualization is an abstract, simplified view of the world that we wish to represent for some purpose." [1, p. 907]. This definition has been criticized on the grounds that it is too vague, and that it has led to the circumstance that a wide range of constructs have become "common bedfellows" under a common rubric [2, p. vi]. Nonetheless, and perhaps because Gruber's definition allows a broad interpretation, it is the definition of conceptualization that is used here. The world that has been captured by these conceptualizations can be as broad as the universe itself, or it can be as narrow as a highly restricted domain, such as the world of restaurant dining [3].

It is the thesis of this paper that it is necessarily the case that *every* conceptualization is biased. This is because representing, or categorizing, the world depends on at least two crucial factors (1) the purpose for which the conceptualization is being created, and (2) the world view of its designer, with the corollary that this depends on the state of general knowledge at the time, as well as on the personal knowledge of the designer. There is bias in the choice of the categories themselves, in the hierarchical arrangement of the categories, in the depth or level of granularity chosen, and in the interrelationships that may or may not be made explicit among the categories. Bias is understood here to be a predisposed tendency toward a certain point of view, which is most often based on a particular system of beliefs. Bias does not immediately make a conceptualization incorrect or nonsensical, nor does it in any way need to diminish the usefulness of the conceptualization. The bias simply needs to be recognized, and its consequences need to be addressed if the conceptualization is to be used for some purpose other than the one for which it was originally designed.

E-mail address: alexa_mccray@hms.harvard.edu.

1532-0464/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jbi.2005.08.007

The paper begins with a brief overview of the disciplines that have been concerned with conceptualizing, or categorizing, particular domains, these disciplines being closely allied with the purposes for which the conceptualizations have been designed and the types of artifacts that result from these efforts. This is followed by a non-exhaustive, illustrative, historical perspective. The paper concludes with the interesting case study of biological taxonomy, which illustrates many of the issues raised in this paper.

1.1. Conceptualizing the world

It is fair to say that all attempts at categorizing the world have as their goal an accurate representation of some reality, in some domain of interest. Philosophers, for example, have and continue to be concerned with fundamental notions of existence. Philosophical ontology, according to one view, is "the science of what is, of the kinds and structures of objects, properties, events, processes and relations in every area of reality." [2, p. iii]. The goal may be to understand the existence of a supreme being, to understand the nature of existence, or, more broadly, to understand the nature of reality. Philosophers since the time of Aristotle have grappled with the relationship between language, thought, and reality, and, in modern times, the views of Frege [4], de Saussure [5], Ogden and Richards [6], and Wittgenstein [7] have been particularly influential. Ogden and Richards, for example, building on the work of both Frege and de Saussure, propose the "semiotic triangle," such that there is a referent (in reality), a symbol (language) that may refer to that referent, and thought (the concept as it exists in the brain). Characterizing how conceptual knowledge is organized in the brain continues to be of interest to both linguists and psychologists. Linguists such as Dornseiff [8] and Trier [9] propose that concepts are grouped together in the brain according to certain shared characteristics, and psychologists in their attempts to understand the nature of cognition, have posited anatomical loci for semantic fields, or domains, in the brain [10]. A New York Times article in the early 1990s reported, "For a century or so, psychologists pondering the brain's memory handling system have suspected that the brain had some system of putting information in categories, with a separate pigeon-hole for categories like dogs, plants, or numbers, each in a separate network of cells...in the past decade a new surge of information has led to the identification of 15–20 different categories that appear to be the brain's own natural categories for knowledge." [11]. Some neuropsychologists have posited a high-level distinction between the semantic domains of natural kinds, such as animals and plants, on the one hand and man-made objects, such as tools and vehicles on the other [12]. Lesions, disease processes and imaging techniques, such as PET scans and functional MRI's are all used as evidence [13]. Recent findings, however, seem to support an attribute-based neural organization of semantic knowledge, rather than basic categories [14].

Closely allied with work in psychology, computer scientists have, over the last forty years or so, attempted to model the world, or some particular portion of that world In some cases, the purpose has been to mimic human cognitive abilities, as in, for example, the CYC project which attempts to capture human common sense reasoning [15]. Often, however, the goal has been to build practical computational systems for applications such as machine translation, information extraction, question answering, and text summarization [16]. The resultant artifacts have been variously called frames, semantic networks, conceptual graphs, and, most recently, ontologies [17–20]. For Gruber an ontology is an "explicit specification of a conceptualization" [1, p. 907], and, similarly, for Poli it is a framework or structure "within which catalogues, taxonomies, terminologies may be given suitable organization." [20, p. 313].

Occasionally, particularly when researchers are working within similar or the same domain there is strong disagreement about what the correct representations should be. Nowhere is this more true than in standardized coding systems. Medical coding systems, for example, have been developed for epidemiological and data comparability purposes, e.g., the International Classification of Diseases of the World Health Organization [21], for billing purposes, e.g., the Current Procedural Terminology of the American Medical Association [22], and yet others have been developed for information indexing and retrieval, most notably the Medical Subject Headings of the National Library of Medicine [23]. Chute alludes to some of the disagreements over the years in the field of medical coding systems: "Many authors have derided one modality or another, failing to recognize that each serves its purpose." [24, p. 301]. Coding systems have in common that they regulate the terminology that is allowed [25]. The NISO thesaurus standard states, for example, that controlled vocabularies for information retrieval must be arranged in a specific order such that "...relationships among terms are clearly displayed and identified... Its purposes are to promote consistency in the indexing of documents, predominantly for post-coordinated information storage and retrieval systems, and to facilitate searching by linking entry terms with descriptors." [26, p. 38]. The standard not only has naming rules, but it also gives guidelines for the hierarchical arrangement of the terms and for allowable inter-term relationships. In spite of the prescriptive nature of these coding systems, there is rarely agreement on the resulting artifacts. There is often dissatisfaction with the concepts chosen, the granularity with which the concepts are represented, and the hierarchical or relational structure into which the concepts have been placed.

Work on biological taxonomies shares some of the same problems. Biologists, in the pursuit of knowledge about living beings, have developed taxonomies to represent biological knowledge as they understand it. Arguments about how to accomplish this have been in evidence since even before the time of Linnaeus. Ereshefsky notes that one area of disagreement is whether a biologic entity should be rep-

resented in a single hierarchical position only (monism), or whether it can be represented in multiple, equally acceptable positions (pluralism). He says, "But why should we prefer a single classification of a multifarious world? Because...taxonomic monism avoids ambiguity, promotes communication, and helps the advancement of science by funneling resources to the most promising project..." [27, p. 41], but later, "Perhaps the world itself, and not our inability to make reliable contact with it, gives us reason to adopt a pluralistic stance toward taxonomy." [27, p. 45].

Topical dictionaries, such as Roget's Thesaurus, that organize their entries within a conceptual framework face similar taxonomic issues. According to Hüllen, these dictionaries are subject to "religious, ideological, political, scientific, or otherwise predetermined world views" of their developers. "Quite often, the authors of these dictionaries seem to be enchanted by the inherent systematicity of their semantic organization and are lured into a perfection and exhaustiveness which overshoots lexicographical purposes." [28, p. 14–15].

1.2. Illustrative historical perspective

All of the disciplines discussed in the previous section have in common that they have created conceptualizations of a particular domain of interest, whether this be for philosophical, scientific, or practical reasons. In what follows, some specific examples from throughout history are presented. Table 1 lists examples of the categorizations that have been developed, together with the individuals who created them and their backgrounds.

The conceptualizations presented in this section clearly have depended on the purpose for which they were created, the training and occupation of the individuals who created them, and the social and scientific milieu in which they were created.

As early as 1200 BC the Egyptian scribe Amenopĕ had as his stated goal to create a universal catalogue that would classify all that was known at the time. His *Onomasticon* (a compilation of names) was intended for "clearing the mind, for instruction of the ignorant and for learning all things that exist: ... heaven with its affairs, earth and what is in it, what the mountains belch forth, what is watered by the flood..." [29, p. 71]. Gardiner, who has undertaken a thorough study of the work states, "Certainly there was never written a book more tedious and less inspired than the Onomasticon of Amenopĕ." [29, p. 24], and "Out of such grotesque beginnings have our encyclopedias arisen." [29, p. 35]. The book clearly did not meet its stated goal, for example, the last of Amenopĕ's categories, "Parts of an ox and kinds of meat," make it seem as though at this point the scribe simply "abruptly abandoned or was released from his tedious task." [29, p. 37]. Nonetheless, the urge to create a universal catalogue of the world can be seen

Table 1
Conceptualizing the world: examples from history (*Major categories are separated by ';'*)

Century	Person/background	Categories
13th BC	Amenopĕ Scribe	Sky, water, earth; Persons, court, offices, occupations; Classes, tribes, types of human beings; Towns of Egypt; Buildings, their parts, types of land; Agricultural land, cereals and their products; Beverages; Parts of an ox and kinds of meat
4th BC	Aristotle Scholar, Philosopher	Substance; Quantity; Quality; Relation; Place, Time; Position; State; Action; Affection
2nd	Julius Pollux Scholar	Gods; People; Relationships; Science and Arts; Hunting; Meals; Trades; Law; Town organization; Utensils
7th	Isidore of Seville Clergyman	Grammar (literature); Rhetoric and dialectic; Four mathematical sciences (arithmetic, geometry, music, and astronomy); Medicine; Law and chronology; Theology; Languages; Alphabetical list of words; Man and monsters; Animals; Universe; Earth; Buildings and fields; Stones and metals; Agriculture; War and amusements; Ships, buildings, garments; Utensils
10th	Aelfric Clergyman	God, heaven, earth, mankind; Parts of human body; Society; Negative features of human character; Intellectual work; Diseases, afflictions, merits; Weather; Birds; Fish; Wild animals; Herbs; Trees; Buildings; War; Human vices
15th	William Caxton Merchant	Formulae for greetings; Objects: house and furniture, food, commerce; Offices, social ranks, names of professions, trades, crafts; Pilgrimage; Counting, money
16th	John Withals Educator	World; Elements: air, water, earth; Man: crafts, housing, city; Society: law, church, family; Life and death: human body, war, senses
16th	Nicodemus Frischlin Poet, Scholar, Educator	God; Nature: (1) universe, space, time, elements, (2) kingdoms of nature; Man: (1) rationality, (2) knowledge and labor, (3) society
17th	Jan Comenius Clergyman, Scholar, Educator	Naturalia, the world as created by God; Artificialia, the world as created by man; Moralia, the way man treats the world with which he is entrusted; Spiritualia, everything concerned with religion
17th	John Wilkins Clergyman, Scholar, Educator	God and the universe; Substance: the four elements, the kingdoms of nature; Accident: quantity, quality, action, relation
19th	Peter Mark Roget Doctor, Scholar	Abstract relations (existence, quantity, time...); Space (dimensions, motion ...); Matter (inorganic, organic); Intellect (formation and communication of ideas); Volition (individual, intersocial); Affections (moral, religious...)
19th	Melvil Dewey Librarian,Entrepreneur	General; Philosophy; Theology; Sociology; Philology; Natural Science; Useful Arts; Fine Arts; Literature; History

in several others over the centuries. For example, Isidore of Seville's work, *Etymologie*, *sive Origines*, written in the early Middle Ages, is a compilation that includes literature, the sciences, medicine, law, history, theology, the universe and the natural world, and the secular activities of man. For Isidore, "the road to knowledge was by way of words" [30, p. 33]. He believed that if he could discover the origins of all words then he would discover the truth about the world. For example, in book 11, Isidore says, "Homo is so named because he is made of humus (earth), as it is told in Genesis..." [30, p. 213]. The work was commissioned by the Bishop of Saragossa, who was interested in a compilation of "all that ought to be known." [30, p. 31]. Brehaut notes, "Although Isidore is not surpassed in comprehensiveness by any one of the line of Roman encyclopedists who preceded him, in the quality of his thought and the extent of his information, he is inferior to them all. Secular knowledge had suffered so much from attrition and decay that it could now be summarized in its entirety by one man..." [30, p. 33]. While the work was disappointing in its lack of insight and low level of scholarly ability, it had a surprisingly strong influence on thought throughout the Middle Ages.

Aristotle's *Categories*, written many centuries earlier, had the philosophical goal of establishing the relationship between language, truth, knowledge, and reality. Modrak says, "The Categories...is a work that claims both to provide a classification scheme for predicates and to articulate the ontological framework that holds true not only of language, but of the world." [31, p. 28]. Rhetoric, or the art of persuasive speech, was of great importance to scholars during Roman times, and, as such, an additional goal for Aristotle was for the *Categories* to be useful for improving public speaking. This goal was shared by Julius Pollux, a professor of rhetoric in Athens, who writing several hundreds of years later than Aristotle, created his own *Onomasticon*. Pollux's categories included words related to gods (e.g., altars, priests, pious and impious people), people and their relationships (e.g., age groups, birth, marriage, masters, and slaves), science and arts (e.g., grammar, rhetoric, dance, medicine, and diseases), law (e.g., court, judges, and trials), and several categories devoted to the daily activities of the time, such as hunting and trading (e.g., dogs, game, buying and selling, games, ladies' toiletry) [28, p. 44]. According to Hüllen, many who developed such onomastica, with their lists of categories and words, were actually interested in representing the world: "...the compilation of such lists can have the aim of incorporating all the knowledge of the world that exists at a given moment, and also the best way of speaking about it." [28, p. 44].

In the 15th and 16th centuries, one important motivation for categorizing all that was known was for the purpose of foreign language learning and teaching. An interesting example of this was the *Dialogues in French and English* written and printed by William Caxton in 1483. Caxton was a British merchant who traded in goods with Belgium and the Netherlands. Some of the goods he imported for the enjoyment of the English nobility were illuminated manuscripts. These manuscripts were most often written in French and Caxton first translated and then printed the manuscripts for his aristocratic clients. "By printing English translations of these works he could corner the market, for he alone would be able to provide members of the English nobility with reasonably priced books which contained what was currently the most fashionable reading matter in their own language." [32, p. viii], In the course of this he created the *Dialogues* which were meant to teach those who were interested in becoming more proficient in French, which had supplanted Latin as the learned language of the day. Note that among categories such as formulae for greetings, offices and social ranks, and household objects, he includes an entire category for counting and money. Withals and Frischlin in the 16th century were also primarily interested in language teaching. Withals work, *A shorte dictionarie for yonge beginners*, was an English-Latin vocabulary meant for young boys, and the categories were chosen such that they could be more easily memorized; "I haue drawen as diligently as I coulde, the propre names of thynges conteyned vnder one kynde, and disposed them in such ordre, that every childe beying able to reade, may with little labour perfitely imprinte them in memory..." [28, p. 175]. Frischlin, a German professor, poet, and teacher, wrote Nomenclator Trilinguis in 1586. The work contained the by now standard major categories of God, nature, and man, and listed all words in three languages, Greek, Latin, and German.

The most sophisticated approach in this tradition was taken by Jan Comenius some fifty years later. Comenius was from the area that is now called the Czech Republic and lived in many parts of Europe throughout his career as an educator. He believed in universal literacy and devoted his life to "the cause of education, both as teacher and educational theorist..." [33, p. 5]. His *Janua linguarum reserata*, published in 1631 and revised in 1657, was meant to teach Latin to school children and as such was a departure from earlier texts that were primarily literary in nature and had little to do with the daily life of the students. [33, p. 16]. According to Hüllen: "Comenius is the probably unique figure for whom a dictionary...was the direct expression of a philosophical and pedagogical idea...the making of dictionaries was directly linked to the elaboration of his system of thought. This system is composed of theological, philosophical, linguistic, and pedagogic elements." [28, p. 371]. Comenius' conviction that reality, thought, and speech were highly interdependent meant that, for him, learning a language involved "...causing one's mind to operate according to the structure of reality." [28, p. 372]. Comenius' four categories, Naturalia, Artificialia, Moralia, and Spiritualia, reflect well his view of the world, which was strongly determined by his religious beliefs.

By the 16th century the predominance of Latin as the language of scholarly communication had begun to wane, and by the 17th there were efforts to develop a universal,

or international, language that would replace Latin [34–36]. Comenius' view was that such a universal language should be based on a description of reality and that this would then also lead to a better understanding of the nature of language itself [36, p. 176]. Bishop John Wilkins in 1668 published his detailed *Essay Towards a Real Character, and a Philosophical Language*, in which he developed a complicated artificial language where ideas were represented by complex signs that denoted the constituent elements of their referents [37]. For example, "...in his word for 'dog' there were symbols to identify the species in general, indicating elements such as 'beast' and 'viviparous' (producing living young) and the species in particular indicating elements such as 'rapacious' (living on prey)..." [36, p. 175]. Sparck Jones points out that although Wilkins' work was a "noble undertaking," it was never used: "...the conceptual classifications defining its content were both far too complicated and too dependent on the particular details of particular scientific, philosophical and religious beliefs." [38, p. 217].

Peter Mark Roget, a physician by training, was also interested, more than 200 years later, in the notion of a universal language. After a long and successful career as a doctor, he began work on his *Thesaurus of English Words and Phrases* in 1848 at the age of 68, completing it six years later. Speaking of his classificatory structure, he says, "Metaphysicians engaged in the more profound investigation of the Philosophy of Language will be materially assisted by having the ground thus prepared for them...such analyses alone can determine the principles on which a strictly *Philosophical Language* might be constructed. The probable result of the construction of such a language would be its eventual adoption by every civilized nation; thus realizing that splendid aspiration of philanthropists, the establishment of a Universal Language." [39, p. xxiii]. His primary aim, however, was much more practical. He was interested in providing a reference work for writers, and particularly for those who were engaged in language translation. He says, "We seek in vain the words we need, and strive ineffectually to devise forms of expression which shall faithfully portray our thoughts and sentiments. The appropriate terms, notwithstanding our utmost efforts, cannot be conjured up at will. Like "spirits from the vasty deep," they come not when we call- It is to those who are thus painfully groping their way and struggling with the difficulties of composition, that this Work professes to hold out a helping hand." [39, p. vi–vii]. But, perhaps, what motivated Roget most strongly was what he saw to be the decline of the English language. In another characteristically hyperbolic statement, he says, "Some modern writers, however, have indulged in a habit of arbitrarily fabricating new words and a new-fangled phraseology, without any necessity, and with manifest injury to the purity of language. This vicious practice, the offspring of indolence or conceit, implies an ignorance or neglect of the riches in which the English language already abounds, and which would have supplied them with words of recognized legitimacy, conveying precisely the same meaning as those they so recklessly coin in the illegal mint of their own fancy." [39, p. xxi–xxii].

Standardizing and regulating library cataloging practices was the goal of Melvil Dewey in establishing his *Decimal Classification*, which was first published in 1876 [40–42]. According to Miksa, Dewey, as well as many others who created classificatory schemes to organize knowledge, believed that hierarchical structures of knowledge categories were somehow natural, "...that is, they reflect the way the human mind produces and uses knowledge and the way such knowledge exists in its various relationships." [42, p. 41]. Dewey, in addition to being a trained librarian — he was a co-founder of the American Library Association—was also an entrepreneur. He worked hard to establish his *Decimal Classification* as the only cataloging system used in U.S. libraries, specifically asking that libraries use it without alteration. When the Library of Congress approached Dewey in 1899 about adopting his system, Dewey refused because the Library intended to make changes to it to serve its own purposes [42, p. 43–44].

1.3. Challenges in representing reality—a case study

The Linnaean system of naming and classifying plants and animals has been used by zoologists and botanists for over 250 years. The taxonomy depends crucially on a system of hierarchical, categorical ranks, or levels. Once a taxon is created, each of its members is required to have the essential properties of that taxon. Linnaeus' original system has five levels, so *Homo sapiens*, for example, belongs to the Kingdom *Animalia*, the Class *Mammalia*, the Order *Primate*, the Genus *Homo*, and the Species *Homo sapiens*. (These five levels have been expanded over the last hundred years so that now there are 21 levels, in the so-called Modern Synthesis [43].) Notice that species names are actually binomials, consisting of the genus and species name together. Linnaeus felt that biologists should memorize all species names, but since he recognized that there were too many of these to make memorization feasible, he reasoned that if the biologist memorized all genus names instead—in his estimation there were some 300 plant genera and just over 300 animal genera—then at least the taxonomic rank of the species would be immediately recognizable. Since, for Linnaeus, the number of genera was set once and for all in God's original creation, the number would not increase and memorization of genus names would be a more reasonable expectation.

For Darwin, whose work appeared more than a century after Linnaeus, and for whom species are the result of gradual evolutionary changes, the appropriate method for classifying organisms was as follows: "All true classification is genealogical; that community of descent is the hidden bond which naturalists have been unconsciously seeking, and not some unknown plan of creation, or enunciation of general propositions; and the mere putting together and separating of objects more or less alike." [Darwin 1859, quoted in [27, p. 29]].

Since the time of Darwin's ground-breaking work, there have been many different schools of thought about how taxonomic work should proceed in practice. One problem is that, as new knowledge comes to light, taxonomies are constantly being revised, such that in certain cases species are moved from one genus to another. Because of the binomial naming scheme, this, of course, means that the name of the species has to be changed as well. Various approaches have been suggested to deal with this problem and other problems inherent in the Linnaean system. The dominant approach to biological taxonomy today is historical; that is, entities are classified according to their causal, genealogical, relations rather than their intrinsic qualitative or essential features. But even within the historical approach there are different ways of proceeding (although most of these are still one variation or another of an adapted Linnaean scheme), and this has led to serious differences of opinion within the community of taxonomists. Ereshefsky says, "The debate among biologists concerning general schools and species concepts has been extensive and at times rancorous." [27, p. 51].

The debate has become even more rancorous in the last few years, with the announcement and development of the PhyloCode system [44–50]. The original developers of the PhyloCode, de Queiroz, and Gauthier, argue for phylogenetic names, stated in terms of common descent and phylogenetic entities [48]. The PhyloCode as it is currently envisioned has seven "Principles," including that names need to be unique and stable, and that relationships are a matter of "hypothesis." Principle six states: "The PhyloCode permits freedom of taxonomic opinion with regard to hypotheses about relationships; it only concerns how names are to be applied within the context of a given phylogenetic hypothesis." [49]. The PhyloCode has not yet gone into operation, but the debates surrounding it are strong. An article in *Science* several years ago reported on interviews with a number of biologists, including Michael Donoghue at Yale and Kevin Nixon at Cornell: "The resulting controversy over the new naming system, known as 'PhyloCode,' has pitted colleague against colleague, office mate against office mate. 'You've got people willing to throw down their lives on both sides,' says Michael Donoghue..." [50, p. 2304], and "...[PhyloCode] is not going to die out, because the spinmeisters behind this have the ear of the large funding agencies...They are going to erect a shadow government and [set up] a coup, Nixon complains. 'This is arrogance'." [50, p. 2307].

2. Conclusion

Early in the twentieth century, the philosopher Hans Vaihinger suggested that the goal of the scientific enterprise is to find the "ultimate truth" [51].[1] Since this is most likely not attainable, he argued that we as scientists should proceed "as if" the constructs we create are true. He refers to these constructs as "useful fictions" since only in this way will science advance: "It is in fact the essential object of science to develop only such ideas as have an objective correlate and to eliminate all admixture of the subjective. Such a task is, however, not easily accomplished...the ideal, in which the world of ideas consists exclusively of congruous, well-ordered and non-contradictory constructs is only to be attained slowly and with difficulty... It must be remembered that the object of the world of ideas as a whole is not the portrayal of reality—this would be an utterly impossible task—but rather to provide us with an instrument for finding our way about more easily in this world." [51, p. 15]. Jacquette recently echoed these ideas, saying that scientific ontologies must be adequate for their individual purposes: "None of these need be *the* ontology in big letters that correctly explains the way things really are..." [52, p. 5]. He goes on to say, "A correct pure philosophical ontology...must aspire to objectivity while remaining open-minded about the existence of subjective phenomena." [52, p. 8].

The philosophers, scientists, and other scholars discussed in this paper have all attempted to create representations that depicted reality as they understood it. They have done this by creating, in most cases, carefully constructed conceptualizations that are intended to represent some domain of interest, including the known world. The resulting artifacts are necessarily biased—where bias is understood to be a predisposed tendency to a certain point of view—because they have been shaped by their designers and the variety of purposes they had in mind. Some have been interested in elucidating fundamental philosophical and linguistic ideas, others in cataloguing all that is known for epistemological reasons, and still others have had more practical concerns, such as pedagogy and standardization for enhanced communication or information access. Depending on our own orientation, we may find some of these artifacts useful and well-formulated, and we may find others amusing, quaint, outdated, ill-formed, or otherwise inadequate.

When evaluating any conceptualization, whether it be historical or one that is created today, it may be useful for us to consider its underlying biases. This also implies that those of us who create conceptualizations must, to the extent we are able to do so, be clear about our own biases. This may mitigate the sometimes acrimonious or even arrogant debates about whether a particular representation is the one and only correct representation in a field of interest.

[1] Thanks to Warner Slack for bringing the work of Hans Vaihinger to my attention.

References

[1] Gruber TR. Toward principles for the design of ontologies used for knowledge sharing. Int J Hum Comput Stud 1995;43:907–28.

[2] Smith B, Welty C. Ontology: towards a new synthesis. In: Welty C, Smith B, editors. Formal ontology in information systems. New York: ACM Press; 2001.

[3] Schank R, Abelson R. Scripts, plans, goals, and understanding. Hillsdale: Erlbaum; 1977.

[4] Frege G. Über Sinn und Bedeutung. Zeitschr f Philosophie und philosoph Kritik 1892;100:25–50.
[5] de Saussure F. Course in general linguistics. New York: McGraw-Hill; 1915.
[6] Ogden CK, Richards IA. The meaning of meaning. London: Routledge & Kegan Paul; 1923.
[7] Wittgenstein L. Philosophical investigations. Oxford: Blackwell; 1953.
[8] Dornseiff F. Der Deutsche Wortschaft nach Sachgruppen. 8th ed, 2004. Berlin: de Gruyter; 1933.
[9] Trier, J. Das Sprachliche Feld. Eine Auseinandersetzung. Neue Jahrbücher für Wissenschaft und Jugendbildung. 1943;10:428–49.
[10] Lichtheim L. On aphasia. Brain 1885;7:433–84.
[11] Hilts PJ. 'Hole' in tumor patient's memory reveals brain's odd filing system. N Y Times Med Sci 1992:C3.
[12] Devlin JT, Russell RP, Davis MH, Price CJ, Moss HE, Fadili MJ, et al. Is there an anatomical basis for category-specificity? Semantic memory studies in PET and fMRI. Neuropsychologia 2002;40:54–75.
[13] Gainotti G. A metanalysis of impaired and spared naming for different categories of knowledge in patients with a visuo-verbal disconnection. Neuropsychologia 2004;42(3):299–319.
[14] Lee AC, Graham KS, Simons JS, Hodges JR, Owen AM, Patterson K. Regional brain activations differ for semantic features but not categories. Neuroreport 2002;13(12):1497–501.
[15] Lenat DB. Cyc: a large-scale investment in knowledge infrastructure. Commun ACM 1995;38(11).
[16] Nirenburg S, Raskin V. Ontological semantics. Cambridge, MA: The MIT Press; 2004.
[17] Sowa JF. Conceptual structures: information processing in mind and machine. Reading: Addison-Wesley; 1984.
[18] Fellbaum C. WordNet: an electronic lexical database. Cambridge, MA: The MIT Press; 1999.
[19] McCray AT. Representing biomedical knowledge in the UMLS Semantic Network. High-Performance Medical Libraries: advances in Information Management for the Virtual Era. Westport: Meckler Publishing; 1993.
[20] Poli R. Ontology for knowledge organization. In: Green R, editor. Knowledge organization and change. Frankfurt: Indeks; 1996. p. 313–9.
[21] International Classification of Diseases. World Health Organization. Available from: http://www.who.int/classifications/icd/en/ Accessed August 26, 2005.
[22] Current Procedural Terminology. American Medical Association. Available from: http://www.ama-assn.org/ama/pub/category/3113.html Accessed August 26, 2005.
[23] Nelson SJ, Johnston D, Humphreys BL. Relationships in medical subject headings. In: Bean CA, Green R, editors. Relationships in the organization of knowledge. New York: Kluwer Academic Publishers; 2001. p. 171–84.
[24] Chute CG. Clinical classification and terminology: some history and current observations. J Am Med Inform Assoc 2000;7:298–303.
[25] Vickery BC. Systematic subject indexing. J Doc 1953:48–57.
[26] Guidelines for the construction, format, and management of monolingual thesauri. ANSI/NISO Z39.19-1993. An American National Standard; developed by the National Information Standards Organization. Bethesda: NISO Press; 1993.
[27] Ereshefsky M. The poverty of the Linnaean hierarchy: a philosophical study of biological taxonomy. Cambridge, UK: Cambridge Universiy Press; 2001.
[28] Hüllen W. English dictionaries 800–1700: the topical tradition. Oxford: Oxford University Press; 1999.
[29] Gardiner AH. Ancient Egyptian onomastica. London: Oxford University Press; 1947.
[30] Brehaut E. An encyclopedist of the Dark Ages, Isidore of Seville. First edition 1912. New York: B. Franklin; 1964.
[31] Modrak DKW. Aristotle's theory of language and meaning. Cambridge, UK: Cambridge University Press; 2001.
[32] Blake NF. Selections from William Caxton. Oxford: Clarendon Press; 1973.
[33] Murphy D. Comenius: a critical reassessment of his life and work. Dublin: Irish Academic Press; 1995.
[34] Clauss S. John Wilkins' essay toward a real character: its place in the seventeenth century episteme. In: Struever N, editor. Language and the history of thought. Rochester: University of Rochester Press; 1995. p. 27–49.
[35] Hüllen W. A history of Roget's thesaurus: origins, development, and design. Oxford: Oxford University Press; 2004.
[36] Subbiondo JL. Universal language schemes and seventeenth-century Britain. In: Koerner EFK, Asher RE, editors. Concise history of the language sciences: from the Sumerians to the Cognitivists. Cambridge, UK: Pergamon; 1995. p. 174–8.
[37] Wilkins J. An essay toward a real character and a philosophical language. London; 1668. (Reprinted at Available from: http://reliant.teknowledge.com/Wilkins. Accessed April 13, 2005.).
[38] Sparck Jones K. Synonymy and semantic classification. Edinburgh: Edinburgh University Press; 1986.
[39] Roget's thesaurus of English words and phrases: facsimile of the first edition 1852. Oxford: Bloomsbury Publishing Limited; 1987.
[40] Dewey Decimal Classification. OCLC Online Computer Library Center. Available from: http://www.oclc.org/dewey/. Accessed August 26, 2005.
[41] Comaromi JP. The foundations of the DDC. In: Stevenson G, Kramer-Greene J, editors. Melvil Dewey, the man and the classification. Albany: Forest Press; 1983.
[42] Miksa FL. The DDC, the universe of knowledge, and the postmodern library. Albany: Forest Press; 1998.
[43] Mayr E. Biological classification: toward a synthesis of opposing methodologies. Science 1981;214(4520):510–6. New Series.
[44] Cantino PD, Bryant HN, De Queiroz K, Donoghue MJ, Eriksson T, Hillis DM, Lee MSY. Species names in phylogenetic nomenclature. Syst Biol 1999;48(4):790–807.
[45] Godfray HCJ. Challenges for taxonomy. Nature 2002;417:17–9.
[46] Thiele K, Yeates D. Tension arises from duality at the heart of taxonomy. Nature 2002;419(6905):337.
[47] Moore G. Down with the Kingdom (phylum, class, and order, too). Science 2002;297:1650–1.
[48] de Queiroz K, Gauthier JA. Phylogeny as a central principle in taxonomy: phylogenetic definitions of taxon names. Syst Zool 1990;39:307–22.
[49] PhyloCode. Available from: http://www.ohiou.edu/phylocode/ Accessed August 26, 2005.
[50] Pinnisi E. Linnaeus' last stand. Science 2001;291:2304–7.
[51] Vaihinger H. 'The philosophy of 'as if', a system of theoretical, practical and religious fictions of mankind. Translated by C.K. Ogden, 1924. London: Routledge & K. Paul; 1968.
[52] Jacquette D. Ontology. Montreal: McGill Queen's University Press; 2002.