On Folk Conceptions of Mind, Agency and Morality

Anna Wierzbicka*

Part A: A comment on the topic of this Special Issue

The formulation of the topic for this special issue is somewhat paradoxical: how can there be different folk conceptions of mind given that 'mind' itself is a folk conception (concept)? Such paradoxical formulations may be convenient for practical reasons; it is important, however, to distinguish convenience from theoretical validity.

To see that 'mind' is a folk concept it is sufficient to note that mind is an English word with no exact semantic equivalent in other languages, or indeed in older English (cf. Wierzbicka 1992). 'Mind' is an important folk concept in modern English, just as 'duša' is an important folk concept in modern Russian, 'kokoro' in modern Japanese, 'maum' in modern Korean, and so on (for detailed discussion, see Wierzbicka 2005, Yoon 2003). All these words imply a dichotomous model of a person, in which a person has two main parts: a visible one (the body), and an invisible one. Linguistic evidence shows that the first, visible part is conceived of in essentially the same way in all cultures, as all languages have a word corresponding in meaning to the English word body (in the relevant sense). The other main part of the person, however, is conceived of differently in different cultures. Roughly speaking, the (contemporary) English word mind presents this part as primarily an organ of thinking and knowing, whereas the Russian word duša, the Japanese word kokoro and the Korean word maum link it in various ways with feeling, wanting, and choosing between what is 'good' and what is 'bad'.

^{*} School of Language Studies, Australian National University. Email: Anna. Wierzbicka@anu.edu.au.

The reason why such different folk concepts established in different languages can be meaningfully compared at all is that there are some universal human concepts identifiable in all languages through distinct words and that these words provide a common "measure" — a necessary "tertium comparationis". As the empirical research carried out within the NSM framework (see below) has established, concepts like PEOPLE, BODY, PART, THINK, KNOW, FEEL, GOOD and BAD belong to the set of sixty or so human concepts for which every language has identifiable lexical exponents (words or word-like elements). These sixty or so universal human concepts provide a "natural semantic metalanguage" (NSM) in which all language-specific folk concepts can be meaningfully compared (cf. Wierzbicka 1996, Goddard 1998, Goddard and Wierzbicka eds. 2002).

Comparing folk concepts through culture-specific English words like *mind* gives such English words an unjustifiable status of neutral analytical tools and ignores the fact that they themselves stand for language- and culture-specific folk concepts which need to be explained and "deconstructed" just like any other folk concepts.

Using NSM, the first question hinted at, in a compressed form, in the title of this special issue can be clarified and sharpened if it is rephrased along the following lines:

Assumptions

- 1. People in all places think about people like this: a person has two parts people can see one of these two parts (the body) people can't see the other part
- 2. When people think about this other part they can think about it in many ways people in one place don't think about it like people in another place think about it

Question

How do people in different places think about this other part of a person?

The two other subdomains identified in the title of this special issue are "agency" and "morality". Unlike *mind*, the word *agency* does not stand

for a folk concept because it is a technical, philosophical term which is not part of everyday language. To be useful as an analytical tool, however, *agency*, too, need to be defined in simple and universal human concepts. Clearly, what is meant by *agency* involves the universal concepts DO, WANT and SOMEONE, but what exact configuration of these three is intended remains to be defined (by anyone who wishes to use this term as a useful analytical tool).

I imagine that the second question hinted at in the title of this special issue is intended to be understood along the following lines:

Assumptions

- People in all places think about people like this:
 people do many things
 people do some things because they want to do these things
 people do some other things not because they want to do these things
- 2. When people think about it they can think about it in many ways; people in one place don't think about it like people in another place think about it

Question

How do people in different places think about it when they think about these things?

(If this is not exactly what is meant by the reference to "agency", NSM can be used to clarify what is actually meant.)

The third term, *morality*, has a somewhat half-way status: it is less colloquial than *mind*, but more so than *agency*. The core concepts on which this semi-technical notion is based are clear enough: they are the universal evaluative concepts good and BAD, as well as the basic predicative concept do. Again, not all languages have a word corresponding in meaning to *morality*, but they all have words corresponding in meaning to *good*, *bad*, and *do*. Thus, the question hinted at in the title of this special issue – what are the different "folk conceptions of morality" operative in different cultures? – can be rephrased and clarified through NSM:

Assumptions

- 1. People in all places think like this: people can do bad things, people can do good things
- 2. When people think about it, they can think about it in many ways people in one place don't think about it like people in another place think about it

Question

How do people in different places think about it?

At least this is how I understand the intention behind the title of this special issue. If my understanding is not quite right, my interpretation can be easily corrected in the same simple and universal semantic metalanguage. (For my own analyses of universal moral beliefs carried out in NSM see Wierzbicka 2001: 161-9.)

These are, then, my main points, which apply, in one way or another, to all the papers in this special issue:

- 1. To compare folk conceptions or folk concepts of any kind we need a tertium comparationis, that is, a culture-independent semantic metalanguage.
- 2. English cannot serve as such a metalanguage, because like any other natural language, it is itself saturated with culture-specific folk conceptions.
- 3. A culture-independent metalanguage in which unbiased comparisons can be carried out is available in "NSM", that is, the Natural Semantic Metalanguage.
- 4. Language is a key issue in all cross-cultural research and all research which has as its subject **human** cognition. No matter how broad the empirical basis of a cross-cultural study, or the study into **human** cognition, is, if this study does not pay attention to the language in which its hypotheses and analyses are formulated, it is likely to impose on the data an ethnocentric perspective. Such ethnocentrism may have been unavoidable in the past, before it was known what the universal, culture-independent human concepts were. Now that this is known, however, it is no longer unavoidable. The Natural Semantic Metalanguage is available as a tested analytical tool for anyone who

would wish to engage in a study of human speech practices, and human cognition, in an unbiased and maximally (if not entirely) culture-independent ways. The effectiveness of this tool has been demonstrated in hundreds of analyses, carried out by many scholars across a broad spectrum of languages, cultures, and conceptual domains (cf. the NSM homepage http://www.une.edu.au/arts/LCL/disciplines/linguistics/nsmpage.htm).

Part B: A comment on Bertram Malle's paper "Intentionality, Morality, and Their Relationship in Human Judgment"

This paper opens with a bold statement about the fundamental categories with which the human mind makes sense of the world:

Considerations of intentionality permeate human social life. In court and sports, in budding relationships and routine interactions, people observe and process behavior and judge it for intentionality. Such judgments are so deeply ingrained in human cognition that we might count intentionality alongside space, time, and causality as one of the fundamental categories with which the mind makes sense of the world.

This opening statement posits four conceptual universals: space, time, causality, and "intentionality". How do these putative universals measure up against the empirical cross-linguistic investigations into conceptual universals? The answer is that the first three are indeed fully corroborated by such investigations: all languages have words (or word-like expressions) corresponding in meaning to the English words when, where and because. As far as the "intentionality" is concerned, however, the situation is not as simple. To begin with, very few languages have a word which could be matched in meaning with the technical English word intentionality, and indeed ordinary English doesn't have such a word either: it does have the words intend, intention, intentional, and intentionally, but not intentionality.

I am not saying that this automatically disproves the claim that "intentionality" is a conceptual universal in some sense, but the question must be asked: in what sense exactly? The claim would be a little clearer if we could assume that what the author really means is what is

known in ordinary English as "intention', but he explicitly denies this ("we contrasted the conditions of ascribing an *intention*...to the conditions of ascribing *intentionality*...).

So what exactly is meant by "intentionality"? The author repeatedly calls "intentionality" a "folk concept", but no ordinary folks have a clue as to what the word *intentionality* means and they don't have it in their vocabulary. We know that the concept of 'mind' is indeed a folk concept shared by the speakers of English, because the word *mind*, commonly used by ordinary speakers of English, provides evidence for it. We also know that concepts like THINK, KNOW, and WANT are parts of the universal "folk model of the person", because empirical evidence strongly suggests that all languages have words for 'thinking', 'knowing', and 'wanting', and also that in all languages these words can be combined with the words for 'someone' and for 'people'. But "intentionality"?

At this point in any discussion about conceptual universals opponents of the NSM theory of language and thought usually make the following move: they question the relevance of words to the study of "human cognition". As one such opponent put it, "let's not deify words" (Lazarus 1995: 259). Thus, it could be argued that all human beings have a concept of 'intentionality' even though they have no word for it.

This is of course not a claim that could be refuted: it is impossible to prove the non-existence of a putative "ens" (in Ockham sense) which has no external manifestations. The burden of proof, however, is not on those who deny the existence of such putative entities (in this case, concepts), but on those who posit their existence. Ockham's razor remains a fundamental methodological principle in the study of any domain, including human cognition: *entia non sunt multiplicanda praeter necessitatem*.

As mentioned earlier, universal human concepts which do manifest themselves in human languages include, inter alia, THINK, WANT, DO and BECAUSE. These concepts enter into certain universally recurring configurations. Sometimes such configurations may in fact correspond to something analogous to what the author has in mind when he uses the word *intentionality*. One suggestion made in his paper is that "intentionality" has four components which can be identified in terms of four English words: *desire*, *belief*, *intention* and *awareness*. All these four words

have complex English-specific meanings, but they all build on certain identifiable universal human concepts or their combinations: in particular, desire includes among its semantic ingredients want and feel, belief, think and know, intention, think and want, and awareness, know and now (for a detailed semantic study of the English words believe and belief, see Wierzbicka, 2006).

I do not know exactly what the author has in mind when he talks about "intentionality", and the enumeration of four semantically complex English words (desire, belief, intention and awareness) does not sufficiently clarify this concept for me. Speaking in ordinary English, one could say that people have "goals", that someone did something "on purpose", or that someone "intends" to do something. Each of these words involves the concept 'want', but the overall configuration of concepts is in each case different. Some conceptual configurations which could be considered in this context can be represented as follows:

- A. this person did something (at one time) because this person wanted to do it
- B. this person thought like this (at one time):
 "I want to do something after some time
 I will do it because of this"
- C. this person did something (at one time)because this person thought like this about something:"I want this thing to happenif I do something, this thing will happenbecause of this, I want to do it"

Roughly, these three scenarios could be linked with the words on purpose, intention, and goal. Is it one of these configurations that the author has in mind when he speaks of "intentionality" as a conceptual universal? If so, then which one? If not, then what other precise configuration? I am not arguing here either for or against any such universal; rather, I'm proposing a metalanguage in which the author could state his hypothesis in a clear, testable, and culture-independent way.

In a final section entitled "Improving methodology" the author writes: "One major problem with virtually all extant studies on morality and intentionality is that participants are forced to make judgments in the researchers' terms, whether on rating scales or forced choice items". This

is an important admission, but the problem goes, I think, deeper than the author seems to realize. Describing human cognition in English (or quasi-English) words like *intentionality*, agency and morality, in itself imposes the researchers' terms on other people's ways of thinking. Using French, Japanese or Esperanto instead of English would of course not solve the problem. What could solve it is to replace the full-blown English, with all its idiosyncratic richness and all the folk concepts embedded in its vocabulary, with a mini-English, an English trimmed to the bone, that is, "NSM English": the English version of the Natural Semantic Metalangauge based on the empirically established set of universal human concepts.

Malle's paper ends with an appeal for a broadly based interdisciplinary approach to the study of human cognition in general and "intentionality and morality" in particular:

The recent explorations into the relationship of morality and intentionality have been compellingly interdisciplinary – with psychologists caring about conceptual problems and philosophers caring about empirical data. I would hope that other disciplines will join this exciting effort. We should invite sociologists . . .; anthropologists . . .; developmentalists . . .; and neuroscientists . . . It will take a while to integrate the various theories and data, but only an interdisciplinary approach is the proper way to understanding this fundamentally human phenomenon in all its complexity.

I was saddened but not surprised to see, once again, the conspicuous absence of linguists from the list of those invited to take part in such an interdisciplinary endeavor. Can any "fundamentally human phenomenon" be studied "in all its complexity" without any attention being paid to language, and to languages?

Part C: A comment on Henry Wellman's and Joan Miller's paper "Developing Conceptions of Responsive Intentional Agents"

Wellman and Miller open their summary of their "key empirical findings" with the following statement: "Normal humans construe themselves and others as responsive intentional agents—persons whose actions and expe-

riences depend on their goals, beliefs, and desires shaped by and responsive to the situations (the physical and especially social situations) in which they find themselves". Looking at this statement from the point of view of cross-linguistic semantics I would suggest that it would be good if it could be rephrased and clarified.

Our main source of information about how "normal human beings construe themselves and others" is language, or rather languages, and linguistic evidence suggests that human beings construe themselves and others as beings capable of 'thinking', 'wanting', 'knowing' and 'feeling'. To begin with, therefore, I would suggest that the words *beliefs* and *desires* in the statement of "key empirical findings" should be replaced with the words *think* and *want*. In making this suggestion, I am not merely being pedantic. Rather, I am drawing on detailed linguistic investigations which show that it is THINK and WANT rather than 'believe' and 'desire' which are universal human concepts; and that the English words *desire* and *believe* stand for language-specific semantic configurations (see Wierzbicka, 2006).

In fact, the words think and want are used repeatedly through Wellman's and Miller's paper, too, alongside believe (or belief) and desire, as if it didn't matter which word is used for describing conceptual universals: think or believe, want or desire. My point is that it does matter a great deal because think does not mean the same as believe, or desire, the same as want, and because it is THINK and WANT, not believe and desire, which are empirically established elements of human construal of people (cf. Goddard and Karlsson 2004; Harkins 1995; Goddard, Forthcoming, Gladkova, Forthcoming).

Linguistic evidence suggests that other important elements of this construal are know, say, feel, see and hear (see Goddard and Wierzbicka eds. 2002). These are all simple (undecomposable) concepts for which all languages have lexical exponents (words and word-like elements). Of course universal human concepts may also include some complex ones, such as 'goals', but these are not necessarily linked with distinct words and their universality may be more difficult to establish.

Can the universality of 'goals' be established on the basis of linguistic evidence? It must be said that while all languages have a word (or another lexical exponent) for WANT, many don't have a word comparable

in meaning to *goal*. At the same time, typologists generally accept that all languages have what linguists call "purposive clauses" or "purposive constructions", comparable to the following ones in English:

She went to the river to bathe. She went to the river for water.

The meaning of such constructions can be represented along the following lines (cf. Wierzbicka 1998):

she did something because she thought like this about something:
"I want to do it
I can do it if I do something else before
because of this, I want to do this other thing now"

Broadly speaking, then, the findings of cross-linguistic semantics are consistent with the "theories of mind" according to which people construe human beings in terms of simple categories like thinking, wanting, knowing and feeling (even though psychologists tend to use instead complex English words like *beliefs*, *desires*, *emotions*, and *awareness*), and also in terms of complex categories comparable to "goals".

But Wellman and Miller go beyond all that and put forward in addition a hypothesis that human beings construe people also as "responsive to the situations (the physical and especially social situations)". What exactly do they mean by that? The term "responsiveness" and the statement that "people's actions are shaped by their goals, beliefs and desires and responsive to the situations" suggest something like a conscious reaction: a person "takes stock" of the situation and does something because of it. This seems to suggest the following idea:

when something happens somewhere, someone in this place can think like this about it:

"I have to do something because of this" because this person thinks like this, this person does something

Is this what the authors mean? As it stands, the remark about human responsiveness to situations is not clear. I suggest that whatever the authors have in mind can be clarified by means of the Natural Semantic Metalanguage. As for the suggestion that "people's *experiences* depend on their goals, beliefs, and desires shaped by and responsive to the situa-

tions" it is even harder to be sure what exactly the authors meant by it, given the complex and culture-specific meaning of the English word *experience* (different, for example, from that of the German words *Erfahrung* and *Erlebnis*). Again, the exact content of the claim could be clarified by means of NSM.

Part D: A comment on Shaun Nichols' paper "Folk intuitions on free will"

The stated goal of Nichols' paper is "to determine the character of folk intuitions surrounding agency and responsibility". In approaching this goal, the author criticizes what he calls "traditional conceptual analysis" (as represented, for example, by Fred Jackson (1998)), for being blind to the distinctive contributions of culture": if one follows this tradition, "one will be unable to recognize which features of one's concepts are culturally local".

I totally agree with these criticisms and I applaud the author for his attention to what he calls "significant intercultural variation in philosophically central intuitions". The question is, however, whether the alternative approach which he advocates and which he calls "empirical conceptual analysis", can escape such criticisms itself.

Nichols suggests that if instead of armchair philosophizing "interviews and anonymous surveys are used", one can "get a more objective measure of the folk intuitions", but he doesn't consider the question of the language in which these interviews and surveys are to be conducted. He doesn't take into account the fact that the English language in which the questions in the interviews and surveys reported in this paper are framed carries with it assumptions which are "culturally local".

Furthermore, in trying to determine "folk views on the nature of agency" (that is, "do the folk treat choice as deterministic or indeterministic") the author apparently asks his interviewees questions which are formulated in such complex philosophical language that even native speakers of English who are "ordinary folk" rather than philosophers are unlikely to understand them. In view of this, it is hardly surprising that the results of the studies reported in the paper are largely negative ("the folk seem to have inconsistent intuitions about agency and responsibility").

I submit that what these studies are in fact testing is not folk intuitions but ordinary folk's ability to make sense of the complex and (to a non-philosopher) bizarre scenarios and questions put to them.

For example, in one pilot study, the author presented 4-5 year-old children "with scenarios of physical events, spontaneous choices, and moral choices". In one of these scenarios, in which "Mary chose to steal a candy bar", the 4-5 year-olds were asked the following two questions: 1. "Okay, now imagine that all that was exactly the same and that what Mary wanted was exactly the same. If everything in the world was the same right up until she chose to steal, did Mary have to choose to steal?". 2. "Okay, now imagine that all of that was exactly the same. If everything in the world was exactly the same right up until the water boiled, did the water have to boil?"

It would seem obvious that these questions (especially the second one) are culturally inappropriate in talking to 4-5 year-old children; and also that they are saturated with "local" cultural assumptions. Suffice it to say that the phrases like "she chose to steal" or "did Mary have to choose to steal?" could not be translated into most languages of the world, including European languages such as my native Polish. The very fact that one can say in ordinary English "she chose to steal" reflects the importance of "personal choice" in modern Anglo culture (again, an English phrase without an idiomatic equivalent in most other languages). English words like *imagine* and *world* don't have equivalents in all other languages either (they do in European languages, but not, for example, in Australian Aboriginal ones); but the idea that someone may "have to choose to steal" is highly culture-specific and it would be difficult to render in most other languages.

The whole thought-world to which the two questions addressed to these 4-5 year-olds belong is an exceedingly "local" thought-world of a modern Anglo-Saxon philosopher and it is an illusion to think that questions of this kind could sufficiently probe folk intuitions of Anglo children, let alone of all children across languages and cultures.

So here is a suggestion: if you want to investigate young children's intuitions about "agency" and "free will", try using NSM (the Natural Semantics Metalanguage, based on simple and universal human concepts). For example, why not ask questions like the following ones:

- When you do something, do you think you do it because you want to do it or because you have to do it?
- Can you do something because you want to do it, not because of anything else?

In fact, I suggest that with adults, too, questions of this kind would work better than scenarios in which people are instructed, for example, to "imagine a universe (Universe A) in which everything that happens is completely caused by whatever happened before it" and then to compare it with another imaginary universe (Universe B) "in which *almost* everything that happens is completely caused by whatever happened before it".

Being used to this kind of brain-teasers, a philosopher may find it difficult to believe that they are unlikely to make sense to many "ordinary folks" and to reveal their "folk intuitions". Rather than arguing about their limited comprehensibility, therefore, I will stress their limited translatability: many languages of the world do not have words for "imagine", "world" or "universe", or expressions which would accurately translate the phrase "completely caused". The words *happen*, *before* and *everything* (= *all things*) would translate well, because they stand for universal human concepts, but the whole sentence is not translatable and it has features which are patently "culturally local".

The question which follows these untranslatable scenarios is this: "Which of these universes do you think is more like ours?". Nichols reports that in this study "the vast majority of subjects answered that the indeterminist universe (Universe B) is most like ours", but presumably most of these subjects (who felt they had sufficiently understood the question to be able to answer it) were educated native speakers of English. To study folk intuitions about "free will" cross-culturally, one could not rely on scenarios and questions of this kind. One could, however, try NSM, along the following lines:

- Can people do something because they want to do it, not because of anything else?
- When you do something, can you think about it like this: "I don't have to do it; I want to do it"?
- Before you do something, can you think about it like this: "I don't have to do this, I can do something else"?

• When you do something, can you know that you do it because you want to do it, not because of anything else?

Questions formulated along these lines could be universally translatable and in principle, universally intelligible.

In his conclusion, Nichols writes: "Experimental techniques and evidence will, I'm confident, help us discover the psychological underpinnings of the folk intuitions that generate philosophical problems". I submit that to discover and to articulate folk intuitions and their psychological underpinnings we need an appropriate metalanguage, different from culture-specific varieties of English (whether in the form of "philosophical English" or even colloquial English, which is also saturated with Anglo culture), and that the Natural Semantic Metalanguage, based on empirically discovered universal human concepts, can be a helpful tool in any such endeavor.

In a recent paper by Nichols and colleagues (Nichols et al. 2004) and entitled "Semantics, cross-cultural style", the authors argue (with special reference to Kripke's theory of names) that "semantic intuitions vary from culture to culture" (p. B1), that "the intuitions philosophers pronounce from their armchairs are likely to be a product of their own culture and their academic training" and that "in order to determine the implicit theories of names across cultures, philosophers need to get out of their armchairs" (p. B9). Semantic intuitions differ indeed from culture to culture, and one major reason for this is that different cultures are linked with different languages. To free oneself from the preconceptions derived from one's own cultural tradition, therefore, it is not enough to "get out of one's armchair", one also needs to get out, to some extent, of one's language. One way to do this is to descend into that deepest core of one's language which is isomorphic with the deepest core of all other languages. As illustrated in this paper, this can be done through the use of the Natural Semantic Metalanguage.

REFERENCES

Gladkova, Anna

Forthcoming Universal and language-specific aspects of "propositional attitudes": Russian vs. English. In Khlentzos, Drew and Schalley, Andrea (eds), Language and Cognition: Interdisciplinary perspectives.

GODDARD, CLIFF

1998 Semantic Analysis: A practical introduction. Oxford: Oxford University

Press.

Forthcoming A culture-neutral metalanguage for mental state concepts.

In Khlentzos, Drew and Schalley, Andrea (eds), Language and Cognition

Interdisciplinary perspectives.

GODDARD, CLIFF AND SUSANNA KARLSSON

2004 Re-thinking 'think': Contrastive semantic of Swedish and English.

Proceedings of the 2003 Conference of the Australian Linguistic

Society. http://www.necastle.edu.au/school/lang-media/news/als 2003/

GODDARD, CLIFF AND ANNA WIERZBICKA EDS.

2002 Meaning and Universal Grammar: Theory and empirical findings. Amsterdam:

John Benjamins.

HARKINS, JEAN

1995 Desire in Language and Thought: A study in cross-cultural semantics. PhD

Thesis, Canberra: The Australian National University.

Jackson, Fred

1998 From Metaphysics to Ethics: A Defence of Conceptual Analysis. Oxford: Oxford

University Press.

LAZARUS, RICHARD

1995 Vexing research problems inherent in cognitive-mediational theories

of emotion - and some solutions. Psychological Inquiry. 6.3:183-196.

Machery, E., Mallon, R, Nichols, S. and S. Stich

2004 Semantics, cross-cultural style. Cognition. 92. B1-B12.

Wierzbicka, Anna

1992 Semantics, Culture and Cognition: Universal human concepts in culture-specific

configurations. New York: Oxford University Press.

1996 Semantics: Primes and Universals. Oxford: Oxford University Press.

1998 Anchoring linguistic typology in universal human concepts. *Linguistic*

Typology 2(3):141-194.

2001 What Did Jesus Mean? Explaining the Sermon on the Mount and the para-

bles in simple and universal human concepts. New York: Oxford University

Press.

Empirical universals of language as a basis for the study of other

human universals and as a tool for exploring cultural differences.

Ethos. 33:2, 256-291.

2006 English: Meaning and Culture. New York: Oxford University Press.

Yoon, Kyung-Joo

2003 Korean maum vs. English heart and mind: Contrastive Semantics of

Cultural Concepts. Proceedings of the 2003 Conference of the

Australian Linguistic Society.