NSM-BASED CULTURAL DICTIONARIES:
FOR LANGUAGE LEARNERS AND BEYOND

Lauren Sadow
Griffith University
l.sadow@griffith.edu.au

Abstract[*]
For language learners, the transition from classroom to immersion is an exhausting and difficult one. Not least because of how language is used differently "in the real world" to how it is taught in classrooms. There are many "insider" dictionaries of language but few dictionaries which take a closer look at the important words and explain them in ways that learners can understand. Natural semantic metalanguage (NSM)'s way of defining culturally important terms and combining them with cultural scripts gives us an opportunity to go beyond the standard realm of definitions and explore the possibilities of what I am calling "cultural dictionaries". This paper will discuss the current opening in learner lexicography to include emic cultural information. It will then discuss how NSM can contribute to such lexicographical practice. Finally, drawing on the first NSM-based cultural dictionary project—the Australian Dictionary of Invisible Culture for Teachers—it provides reflections, advice, and recommendations for future cultural dictionary projects.
Keywords: learner lexicography, minimal languages, user needs research, cultural semantics, e-lexicography, dictionary-making

1. Introduction

For major world languages, there are many "insider" dictionaries of how to use the language—Australian English, for example, has a plethora of guidebooks on how "native" speakers use English, full of slang, casual and informal speech, and profanities—but there are few dictionaries which take a close look at culturally significant words and concepts and explain them in ways that learners can understand. There are even fewer dictionaries which include descriptions of speaker intentions and attitudes. The manner of defining culturally important terms and combining them with cultural scripts through the natural semantic metalanguage (henceforth NSM) approach gives us an opportunity to go beyond the standard realm of definitions in

---

lexicography and explore the possibilities of what I am calling "cultural dictionaries".

Cultural dictionaries are, broadly speaking, dictionaries which describe culture (also broadly defined) either as the focus of the content, or in addition to more traditional lexicographical content. Past cultural dictionaries fall into other sub-categories of lexicography such as encyclopaedic dictionaries, specialized dictionaries, learner dictionaries, and so on. In this paper, I will focus on NSM-based cultural dictionaries—which use an NSM approach to culture and cultural concepts. Cultural dictionaries as conceptualized in this chapter are relevant to all languages, but the examples here are focused on English.

As discussed in the introduction to this issue, the NSM approach is founded on the idea that there is a consistent core of concepts across all languages, and we can leverage this core to communicate and explain ideas across languages without enforcing a particular linguistic framework on those ideas (Wierzbicka and Goddard 2013). Over more than 40 years of research, proponents of the NSM approach have established 65 semantic primes as that core, plus a variety of lexical and semantic molecules which are very common (near universal) (Goddard 2021).

One of the stand-out unique features of the NSM approach is that it enables descriptions of both meaning and culture at the same time. Moreover, the expressions used in the NSM approach are such that even though the depth and amount of information is great, the language simplifies it so that it can be accessed by people at various levels of linguistic ability.

The 65 semantic primes of NSM can be found in all languages, and semantic molecules have equivalents or close equivalents in most languages on earth. This fact alone means that explications written in NSM are more cross-translatable than other kinds of definitions, making them ideal for language learners. NSM explications (explanations of a word's meaning written using primes and molecules) not only capture the *meaning* of a word, but also its cognitive and affective significance to speakers of a language—how they *think* when they say this word, what they *feel* when they use this word, and what they *want* other people to think and feel when they hear this word. In traditional lexicography this information is rarely

captured outside of labels on entries (like "derogatory" or "informal"), and the cognitive processes are almost never explained.

In this paper, I will first explain the concept of a "cultural dictionary" and elaborate on the kind of information which should be included in one, in comparison to other lexicographical projects. Following this, I will discuss how the NSM approach can be used in lexicography, focusing on the idea of reductive paraphrase, and previous NSM-based dictionaries, including discussing the advantages of a "cultural dictionary" that incorporates this information. Finally, I will dedicate the last section to advice on conducting an NSM-based cultural dictionary project, based on the work done for the *Australian dictionary of invisible culture for teachers* (AusDICT).

2. What are cultural dictionaries?

Lexicography and lexicographers espouse a broad range of types of dictionaries, not all of which conform to the "prototypical dictionary" such as the Oxford English Dictionary. In the middle of the lexicon–encyclopaedia debate (Haiman 1980; Peeters 2000b; Silverstein 2006; Sánchez 2010) is the position that information about words themselves, and how they are used, belong to lexical information and not encyclopaedic. Therefore, the form which a "dictionary" can take is theoretically very broad.

2.1. Encyclopaedic and specialized dictionaries vs cultural dictionaries

There are many contested distinctions between dictionaries and encyclopaedias. By one definition, a "dictionary" should only contain semantic information and any other information is the purview of another discipline (Peeters 2000a). In an alternate definition, dictionaries contain information about a topic, regardless of the kind of information or the kind of topic. The heart of this debate about the difference between dictionaries and encyclopaedias is the distinction between linguistic and encyclopaedic knowledge. While some researchers argue that there is no theoretical distinction between the two, many argue that there is one (see discussion in Peeters 2000b). Even where there is agreement that there is a border between the two kinds of knowledge, the exact location of that border is difficult to determine. As Peeters comments in his introduction to the 2000b volume "the distinction, when made, is not being made along the same lines by everyone with an interest in the matter" (2000a:2). From an NSM linguistics perspective, Goddard (2011) draws a distinction

between the two by saying that "linguistic knowledge is essentially shared between all the speakers of a language, whereas real-world knowledge is not" (Goddard 2011:16). He then continues to specify that folk knowledge is included in linguistic knowledge, as it too is shared among all (or almost all) speakers of the language. In this distinction, scientific knowledge or technical knowledge is limited to encyclopaedic information. On the lexicography side, Silverstein (2006) also makes the point that lexicographical works are, in part, ethnographic and should also reflect uses of language. Based on these discussions, we can consider all information about language which is shared by most speakers to be linguistic knowledge, and therefore can be included in dictionaries—if there is a distinction between encyclopaedias and dictionaries at all.

Specialized dictionaries are defined as those which address specific material, usually in a particular domain (Landau 2001). These dictionaries are especially useful in domains such as business or sciences and can be extensively used in English for Specific Purposes courses (Fuertes-Olivera 2010). Because specialized dictionaries are focused in a way that general dictionaries cannot be, they usually contain fewer entries, and also more information about context and usage than a general dictionary is able to give. In many cases, specialized dictionaries contain entries which do not occur in general dictionaries, especially not concise or learner's dictionaries, as these are dedicated to words which are frequent and necessary to communication.

It would be a reasonable conclusion, then, to say that cultural dictionaries are in fact specialized dictionaries, and ones which specialize in culture. One useful model of thinking about culture is the division of invisible vs visible culture and big "C" Culture vs little "c" culture (see Figure 1). Of course, this division is artificial, but helps to describe the different approaches to culture in both previous "cultural dictionaries" and NSM-based ones. The type of cultural dictionary which I argue for in this paper is one which prioritizes big "C" invisible culture (although the other divisions are not excluded).

	Big 'C' Culture *Classic or grand themes*	**Little 'c' culture** *Minor or common themes*
Invisible culture *"Bottom of the iceberg"*	**Examples:** Core values, attitudes or beliefs, society's norms, legal foundations, history, cognitive processes	**Examples:** Popular issues, opinions, viewpoints, preferences or tastes, certain knowledge (trivia or facts)
Visible culture *"Tip of the iceberg"*	**Examples:** Architecture, geography, classic literature, presidents or political figures, classical music	**Examples:** Gestures, body posture, use of space, clothing style, food, hobbies, music, artwork

Figure 1: Peterson's (2004:25) illustration of the different elements of visible and invisible culture.

This perhaps leaves the question not of whether a cultural dictionary *can* be called a dictionary, but whether it *should* be named as such. I would argue that the name of "dictionary" is a more accurate description for users of what such a work would contain and how it should be used than other more accurate, but perhaps less transparent names (such as "lexicographical resource").

2.2. Culture in other dictionaries

Culture has not been ignored in lexicographical practice, but nor has it been standardised. Many dictionaries which include "culture" in the title, such as *The new dictionary of cultural literacy* (NDCL) (Hirsch et al. 2002) focus on visible culture—especially on people, places, and events—and do not capture the behavioural implications of having this knowledge. The function of such dictionaries is primarily to fill in the "prior text" of visible Culture for users. Other dictionaries which give special attention to culture (noted in their introductions for example) primarily give semantic information but supplement it with cultural context in the form of cross-referencing to larger cultural concepts. One example of this is in the *Dictionary of Hong Kong English* (DHKE) (Cummings & Wolf 2011). These dictionaries aim to draw connections between language and culture, in terms of cultural practices. The DHKE uses frameworks from cultural linguistics

(Sharifian 2011, 2014) to achieve its goal, particularly applying the concepts of cultural schemas and cultural conceptualizations.

Other dictionaries focus on semantic content but make conscious, culturally relevant choices in terms of what entries to include. Among the thousands of commercially available slang dictionaries, one more culturally focused example of this for Australian English was the *Australian cultural dictionary* (AusCD) (Miller et al. n.d.), sadly no longer available. The headwords in the dictionary were limited to Australian slang and artefacts but were chosen for their relevance to the users of the dictionary (migrants to Adelaide) and their known struggles of understanding. Some of the definitions give additional information to the basic semantics, like a small plain-language history of specific foods, somewhat like the NDCL mentioned above. Many of the slang terms in dictionaries such as these are emblematic of cultural attitudes.

Another way in which culture is presented in dictionaries is in a kind of descriptive dictionary with unusually long entries. In these dictionaries, the entries unpack cultural backgrounds based on a word or even a nationality. One example of this, produced by a migrant resource centre in Canberra, Australia, is the *Cultural dictionary and directory: Of people from culturally and linguistically diverse backgrounds* (CDD) (Khalidi et al. 2012) which provides the user (people working with migrants and refugees) paragraphs of information about history and changes influencing particular parts of culture. In some cases, they even describe some of the behaviours and attitudes influenced by that part of the culture. Dictionaries in this style are aimed at a monolingual audience who interacts with many different cultures. It is interesting to note that the CDD, while providing the most information (out of these examples) on invisible culture, still does not provide any suggestions for having successful interactions with people from different backgrounds, and nothing which can be implemented by the reader.

Finally, it is worth noting the Longman Essential Activator (LEA), which is a product of one of the major ESL dictionary publishers. While the LEA does not advertise itself as containing "culture", it does have a section on interaction in the centre, demarcated by different colour pages. This feature is unique to this book (and doesn't appear in other Longman publications) but implicitly teaches invisible culture through teaching interactional

routines, such as accepting an invitation. The disadvantage of this approach is that it teaches specific routines, rather than guiding principles which can be applied to many different kinds of routines and interlocutor goals.

But what is missing from these five approaches is one which addresses invisible culture, not just as an overall summary of the apparently important features, or as a set of specific routines, but as the values and guiding principles which influence interaction and performance of identity and culture.

3. How does NSM fit in?

The obvious parallels between the work of writing NSM explications and that of lexicographers has never been ignored in NSM research. Throughout the history of NSM, there have been a number of discussions and debates pertaining to the connections between it and the field of lexicography (see Wierzbicka 1992, 1996; Goddard 2017b). NSM has been developed in part as a response to the tendency in lexical semantics to capture technical meaning but not folk knowledge—this includes definitions in lexicography. In particular, Wierzbicka (e.g. 1985, 1987) saw that dictionary definitions were often circular and did not follow "the golden rule of lexicography" (see e.g. Atkins & Rundell 2008)—that definitions should be defined in terms simpler than the term being defined. Determining a metalanguage of simplest possible terms for use in definitions resolves this, in turn resolving the problem of circularity. However, because the semantic primes are the simplest level of meaning, it is impossible to define them in simpler terms. Wierzbicka (1996) responds to this by saying that as the concepts are the semantic core of all languages, they should not need definition at any point. In reality, however, dictionaries still include entries for primes (e.g. "think" and "know") albeit often fairly obscure. In addition to the problem of circularity, inaccuracy and obscurity (see Goddard 2011) are two other commonly stated problems in lexicography. NSM explications avoid inaccuracy by ensuring they predict the range of usage, but do not over-predict. In terms of obscurity, explications aim to ensure maximum clarity by using clear and intelligible language in a literal way.

Goddard (2017b) expands on this idea of accuracy by presenting a number of versions of explications using NSM and minimal English (see introduction to this special issue and Goddard 2017a, 2021; Sadow 2020b), with longer definitions (closer in style to NSM)

considered more accurate and shorter definitions (closer to minimal English) to be of "medium accuracy".

There are two kinds of responses which lexicographers are likely to make to these criticisms of circularity, inaccuracy, and obscurity. First, the space requirements of dictionaries, and the need for definitions to be as concise as possible means that NSM compositions are difficult to incorporate into traditional dictionaries. Goddard's (2017b) discussion of concise and precise definitions in minimal English goes some way to ameliorating this critique. Second, the information needs and user needs of dictionaries are not necessarily compatible with the ways in which NSM explications can be seen to over-provide information (Atkins 2008; also see discussion of Barrios Rodríguez's work 2013, 2020 and in section 3.3.3). Lew and de Schryver (2014) point out that many users want very specific information from their dictionaries; and Atkins and Varantola (2008) find that dictionary users are not often willing to read a whole article to find the information they want, meaning that maximally detailed and accurate NSM compositions are potentially too long for dictionary users (see also Lew 2015). For example, the NSM explication of *dogs* (Goddard 2018) is over 900 words long, while the Cambridge Dictionary definition is just 19 words (see Appendix for full explication and definition).

3.1. NSM as a limited defining vocabulary and reductive paraphrase

The main advantages of using NSM as a method of cultural description in language teaching are as follows. First, the NSM approach explains meaning through extensive research into concepts—in the form of words and phrases, or values and attitudes. The meanings are not kept isolated from the pragmatic or cultural implications of the concept, and therefore an NSM composition can provide a better representation of the socially constructed meaning of concepts.

Second, using a limited set of words as a defining vocabulary, based on universality and translatability, guarantees that the resulting definitions are expressed in terms simpler than the ones being defined. This ensures that the entries are both non-circular and easy to understand. Using restrained defining vocabularies is not new to lexicography, but the core differences with NSM are that the defining vocabulary is both much smaller than others (such as the Longman

Defining Vocabulary at 2,000 words (Summers 2006)), and that it is carefully constructed for cross-translatability, rather than being based on word frequency in English.

Third, compositions are presented from an insiders' perspective. That is to say, they generally contain components which describe a way of thinking "I think like this" or "people can think about it like this" with the thought pattern then represented as direct thoughts. They are not abstract concepts or impersonal definitions, they capture how an individual can think and feel, including about concrete objects.

Fourth, because the subset of words and their related grammar are central to all languages, this ensures that definitions can cross the boundaries of languages and cultures. This results in people learning, or unfamiliar with, a language being able to understand the cultural insider's perspective, despite having limited or no prior knowledge of the language or culture.

Fifth, because compositions have a specific structure in addition to the limited language, using the NSM approach levels the playing field between languages as it does not prioritize concepts from one language over another (e.g. prioritizing English-speaking ways of thinking and organizing information). Compositions are structured and presented similarly regardless of language or concept, resulting in comparable entries in a dictionary. In other words, the NSM approach can defamiliarize familiar concepts, making them appear equal to unfamiliar concepts. Thus, the approach allows students and teachers to critically engage with invisible culture across language boundaries. The following examples of pedagogical scripts from Peeters (2019:183–85) illustrate how, although all slightly different, the expressed attitudes can be discussed more easily by using the scripts, rather than by using the headings.

An Anglo pedagogical script connected with epistemic reserve and openness

in Britain, when you say about something: "I think about it like this", it is good to
  say at the same time:
      "I don't say that I know this
      I know that someone else can think differently (from me)."

A French pedagogical script connected with the propensity to take a stand

in France, it is good to say about many things: "when I think about it,
  I think like this: […]"
  you can say it when you know a lot about these things
  you can say it when you know little about these things
  you can say it before you know anything about these things.

A Russian cultural script connected with *vyrazitel'nost'* 'expressiveness'

people think like this:
  "when someone is with other people, it is often good if this someone thinks
  like this:
    I want these people to know what I think,
    I want them to know what I feel."

Finally, concepts are explored in compositions by breaking down the whole into individual components. This allows us to see how each component functions within a concept, and where those components are present in other concepts. Once a concept is deconstructed, it can be reconstructed through each of these components, and seen as the centre where all these components overlap. In the above examples, the phrase "I think like this" or similar is repeated in slightly different framings. By looking closely at where and how these are expressed in the scripts, a learner can develop an understanding of different ways of expressing opinions in those languages and in the world at large.

3.2. Cultural norms and values in addition to language
A well-designed NSM-based cultural dictionary can connect the vocabulary and speech routines which are traditionally learned in language courses to norms of interaction which can guide learners through unfamiliar contexts and help them to interpret speaker meaning in conversations. Outside of language learning, they can also help speakers of a language uncover and reflect on their own speech practices and develop empathy and understanding for other norms, ways of speaking, and ways of expressing oneself.

While many ethnographic studies capture the same information, the NSM approach can strengthen ethnographic research by distilling it into clear and explicit language, while maintaining an insider's perspective. Cultural scripts (see introduction to this volume) written

using NSM which describe norms and values have the same structures and formats as explications, which means that dictionaries can have consistency across entries, even if the entries are of different kinds. Cultural scripts are also easily able to be linked with explications of words. These links then create a network of concepts which criss-cross and interact with one another, providing users with a more intricate and deep understanding of language and its use.

3.3. What NSM-based dictionary projects have been done?
There have been four lexicographical projects based in NSM at the time of writing. However, only one of these (AusDICT, see below) is designed with the intention of foregrounding culture by using explications and cultural scripts. As discussed earlier, one of the strengths of NSM is that explications combine semantic and conceptual meaning in one place, so all lexicographical projects using NSM are cultural to some degree, even if it is not their focus. This section will give a brief overview of these four projects, some of their features, and their relationship to NSM-based cultural dictionaries.

3.3.1. *English speech act verbs*
The first dictionary using NSM was Wierzbicka's *English speech act verbs—a semantic dictionary* (1987). Even now, it is the only traditionally published dictionary of explications, although the field of NSM has changed dramatically since publication. The dictionary is designed for two purposes—to "be of service to the general public—both to native speakers of English and to people learning or teaching English as a second language" and "to be a study of an important section of the English vocabulary" (Wierzbicka 1987:1).

The main body of the dictionary is organized through 37 semantic categories, referred to as "groups". Overall, there are approximately 280 items in the dictionary. Each explication has a page or more in the dictionary, although they are not set out with a single page each. As Wierzbicka (ibid.) notes, this length of explication is much longer than a standard definition in a dictionary, especially since each headword captures only a single meaning (polysemy is dealt with through individual headwords for each meaning).

As implied by the title, this dictionary does not cover the whole of the English language, only speech act verbs. At the time when it was written, cultural scripts were not prominent in NSM literature, and as such the interactional norms and communicative strategies are

not included in the explications. That said, the explications do elaborate on many cultural ideas and norms associated with each speech act verb which make them more useful to language learners than traditional definitions. For example the explication for "thank" (Wierzbicka 1987:214) contains the component "I say this because I want to cause you to know what I feel towards you" which is additional pragmatic information for learners that indicates that it is the spoken thanks which is how the appreciation is expressed, rather than just by actions.

3.3.2. *Learn these words first*

A second example of a dictionary using NSM as a theoretical base is the online educational dictionary called *Learn these words first* (LTWF) (Bullock 2014a, 2021). The research for this dictionary was an original project to create a non-circular dictionary bridging the NSM set of 61 primes (at the time) and the *Longman dictionary of contemporary English*—specifically the Longman defining vocabulary (Bullock 2011, 2014b). The LTWF dictionary is a learners' dictionary of English based on a multi-level structure, presented as a lesson series that builds a learner's vocabulary to the 2,000 most common words in English.

While not technically written in NSM, as the definitions are not explications, nor are they designed to capture every part of the meaning of the word, the non-circular nature of this dictionary is an excellent example of how NSM can be used to build vocabulary for language learners. Most of the definitions in this dictionary could more accurately be described as being in minimal English since they use semantic molecules as the second layer of building blocks to define more complex words, as well as a more idiomatic way of expression. Although this reduces the translatability of the entries, they are still built up from primes, so as students acquire more of the language, they require less translatability. However, the definitions given are much more like traditional dictionary definitions and give no indication of the connected cultural concepts underpinning each of the words, nor do they provide any information about interaction or usage.

3.3.3. Minimal and inverse definitions

Barrios Rodríguez's work on minimal and inverse definitions (2013, 2020) has shown that while explications are excellent tools for defining deep cultural meanings, for many words, fluent speakers do not need such in-depth definitions to identify a word based on a

minimal language composition. Her work focuses on what she calls "basic nouns" such as selected animals, plants, and foods, and looks closely at what kinds of definitions can be considered "accurate" in that they point at the correct referent.

The dictionary project which Barrios Rodríguez discusses in her 2020 paper is not trying to be a cultural dictionary; rather, its emphasis is on providing different levels of definitions for different kinds of users. However, it still demonstrates the power of using NSM as a defining vocabulary in dictionary definitions and shows how NSM and minimal languages can be adapted for different dictionary purposes and users.

3.3.4. *The Australian dictionary of invisible culture for teachers*

The Australian dictionary of invisible culture for teachers (AusDICT) (Sadow 2020a, ausdict.translatableenglish.com) is the first NSM-based cultural dictionary of the kind described in this paper. It focuses not only on words and slang, but also the value and attitudes which underpin their usage. The AusDICT is an online dictionary, aimed at English language teachers in Australia. It contains 333 entries which are grouped into 12 topics, which themselves are divided into further sub-topics.

This dictionary is the only current one which includes both explications and cultural scripts alongside one another. It also is the only NSM-based dictionary to connect concepts to one another across the bounds of the initial topics and sub-topics. Unusually, it also includes a "part of speech" categorisation which subdivides the entries by both the usual part of speech categories (noun, verb, etc.) as well as ones specific to culture (value, attitudes, phrases).

This dictionary uses a variant of minimal English for its definitions, standard translatable English (STE) (Sadow 2021). STE is a minimal English which is designed for applied contexts, with additional molecules carefully selected for the contexts and the topics included. The formatting of STE for this project was guided by user needs research (see section 4.1), and in the final iteration, preferred by the users, was much closer to more traditional NSM explications and cultural scripts than initially expected.

As this dictionary is the only completed NSM-based cultural dictionary project at the time of writing, the remainder of this paper will make reference to the creation process of the AusDICT.

4. How do we make cultural dictionaries?

The following section outlines some of the key considerations for creating an NSM-based cultural dictionary, drawing on my experience creating the AusDICT. It is important to note that these are personal observations from one single project and different projects with different scopes and different users will face their own successes and pitfalls.

In general, the process for creating a cultural dictionary is much like any other lexicographical project, but in some ways more complex as the entries are longer, the scope is broader, and there are fewer established norms and practices.

4.1. User needs research

While it is possible to create a dictionary of any kind without user needs research, as Atkins and Rundell (2008) and Landau (2001) make very clear, best practice is to conduct extensive user needs research before embarking on the production of a dictionary. User needs research helps the lexicographer to design and create a dictionary which people will turn to. If one is looking into traditional publication avenues, then this will overlap with market research.

Landau (2001) describes several different types of information which need to be uncovered during such an analysis. While the exact types of information differ from dictionary to dictionary, they can be broadly grouped into three categories: a description of the target audience; why the project will benefit from knowing their needs; and an understanding of what questions the users are trying to answer (when they use the new dictionary).

Perhaps the most important question to ask in user needs research is that last statement: what questions are users trying to answer by using the new dictionary? Why aren't they using a traditional dictionary? Most of the previous NSM-based dictionary projects which I discussed in section 3 are intended to help users answer very different questions to an NSM-based *cultural* dictionary, and it is important to keep this in mind. For example, a user consulting Wierzbicka's *English speech act verbs* is looking for the fine-grained distinctions between many specific words and is asking a question like "What word expresses what I want to say best?". Whereas a user of an NSM-based cultural dictionary is more likely to be looking for information crossing several genres and ways of speaking such as

"How do I explain my concerns to the teacher?" or "What caused this miscommunication today?".

There are many ways to gather user needs information, but common ways are through surveys, interviews, and focus groups, or even a combination of approaches (Landau 2001; Atkins and Rundell 2008). For the AusDICT for example, I used surveys and focus groups to do user needs research and user testing. User needs research uncovered detailed information about my target users' (English language teachers in Australia) educational backgrounds, perspectives on language and culture, use of dictionaries and textbooks in language classrooms, practical considerations such as technological access during teaching hours, information gaps which they wanted filled, and even design and access preferences.

User testing was an additional component which I undertook to verify how I had implemented the user needs analysis. The user testing took place in two phases, both with focus groups, but the second time participants also completed an evaluation survey, which was based on the stated desires from the user needs, plus the design goals of the project. Both the user needs research and the user testing uncovered perspectives and reactions which I would not have considered without that user input.

4.2. Scope and content selection
The scope of a dictionary project is inherently specific to the project itself, but there are some elements which are worth considering.

In the past, one of the major barriers to using NSM in lexicography has been the limitations of printed dictionaries and hard copy material. With limited space available, lexicographers prefer short and condensed definitions. But with the progress of technology, it is more and more practical to produce digital resources which contain many entries where length is unconstrained. That being said, many mainstream dictionaries have thirty to forty thousand headwords, and it is impractical to produce so many NSM explications or minimal language compositions from scratch (despite how much we wish we could).

My recommendation for new cultural dictionary projects is to constrain the scope of the dictionary in such a way that it can be easily expanded upon with subsequent research. One way of doing this is by choosing specific themes, topics, or domains (as Wierzbicka (1987)

did with speech act verbs). Another would be through selecting particular scenarios such as job interviews, or interactions with doctors. The AusDICT contains 333 entries in total, across 12 topics. The topics were chosen based on priorities from the user needs surveys and previous research, while others that did not make the list have been prioritized for future research.

Another approach to content selection is to think about existing scripts and explications and aiming to fill out interconnected networks based on these concepts. The stratification of cultural scripts into levels—master-level, high-level, mid-level, and low-level/interactional-level (see Sadow 2018)—can help to create these networks and connect concepts to one another—I will return to connecting concepts and cross-referencing shortly.

4.3. Content research—using existing research and doing new research

Research for a cultural dictionary project contains several (overlapping) phases. The first phase is the user needs research as discussed above, which will help to determine content, scope, and features. The second phase is semantic and cultural data collection to write cultural scripts and explications for entries. The third phase is writing and refining the entries, and the fourth is again user needs research and feedback focused on design and format.

The time investment of writing NSM explications and minimal language compositions is a barrier to NSM-based cultural dictionary projects which should not be ignored. Barrios Rodríguez has tried to ameliorate this issue by using students to help write minimal inverse definitions for her project (2020). However, this is not always possible, or practical. Lexicographers writing traditional dictionaries have many variations, editions, and examples of definitions to draw on. Drawing on an existing base of NSM definitions is the easiest way to begin building an entry list for a cultural dictionary, but for most languages there are insufficient explications for a satisfying lexicographical project.

As such, conducting original research is an inevitable part of a cultural dictionary project. To this end, I recommend clearly defining the scope of the dictionary so that research can be consolidated into scripts and explications as efficiently as possible. For the AusDICT project, although I had a large database of explications about English, and Australian English, more than 110 original explications needed to

be written to cover priority areas or to fill in conceptual gaps between high-level cultural scripts and detail-oriented interactional scripts. As discussed above, choosing content around themes can help to focus the research needed, and allow for best-practice semantic, ethnographic, and cultural research to be undertaken.

4.4. Design features

An important part of producing a dictionary is considering the design elements, both in terms of the information to be included in each entry, and in terms of how that information will be formatted and styled. The AusDICT went through a lengthy user needs and creation process which resulted in several iterations of the dictionary, both as an ebook and as a website. I aimed to blend new and traditional elements of dictionary structure and design, so that familiar elements would be intuitively understood, while unfamiliar elements were clear and unambiguous (following Lew 2015). It is unnecessary to go through all the possible elements of a dictionary entry here, Landau (2001) and Atkins and Rundell (2008) and others cover this in far more depth than I could. However, I will mention some elements which both the user needs analysis and the project itself revealed as particularly salient to an NSM-based cultural dictionary. For the most part, I will discuss the AusDICT in its current form as a website, although some reference will be made to the first ebook format. As a reference, I have included an example in Figure 2, with the different parts I will discuss enumerated.
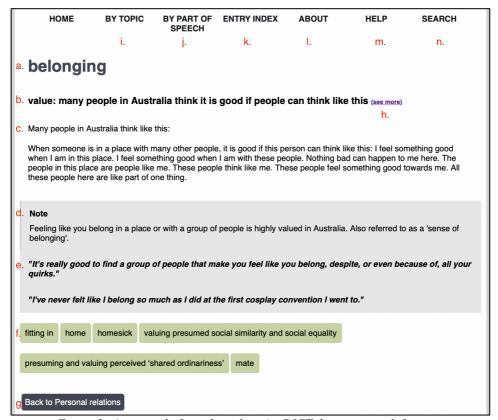
Figure 2: An example from the online AusDICT dictionary with features enumerated. a. headword; b. 'part of speech'; c. entry; d. note; e. examples; f. cross-references; g. and h. links to other groupings; i. j. k. n. navigation pathways; l. and m. 'frontmatter'.

4.4.1. Frontmatter

Because an NSM-based cultural dictionary is something that comparatively few people have been exposed to, explanatory material in the dictionary became essential. In traditional lexicography, this is called 'frontmatter' and, as expected, is at the front of the physical book. In an ebook, it can be literally at the front, but hyperlinked from anywhere, and in a web-based dictionary, it is its own page. The frontmatter in the AusDICT provides two functions—to explain NSM and minimal English, and to explain how to use the dictionary and what a user should be able to get out of it. In the ebook version, this was provided first as the fully described version, and then subsequently as an abbreviated illustration. The abbreviated illustration was then hyperlinked from each entry as a "help" button. In the web version, the two parts are separate pages, accessible

through the main menu which is always visible to users (in Figure 2, l. and m.).

4.4.2. "Parts of speech"

Traditional dictionaries usually contain a label in each entry which specifies the part of speech of the entry. This is good for users because it disambiguates the meaning which is defined and provides some usage context for the definitions. For words, this is easily done, but for cultural scripts, values, phrases, and other kinds of NSM compositions, this type of context is still useful but more difficult to specify.

For the AusDICT, I decided to extend the traditional "part of speech" label to include labels for all of these types of entries. I also decided to extend each label by applying a minimal English version of the label in addition to the linguistic term (e.g. Machin 2021). As a result, some of the more traditional labels were subdivided into more specific groupings (e.g. *adjective* had three variants "adjective: a kind of something", "adjective: someone feels something", and "adjective: a kind of person"). Some examples of the non-traditional additions to these labels were "phrase: someone says these words" and "value: many people in Australia think it is good if people can think like this" (as in Figure 2, b.). The purpose of these labels was to help provide context (as previously mentioned) but also to help users navigate through entries.

4.4.3. Language examples

Although not standard in dictionaries, examples are common in various types of dictionaries, for various purposes. For the purpose that the AusDICT was intended, the user needs research revealed that examples would be useful on each entry, especially for designing lessons. As a result, each entry is accompanied by several examples which illustrate the usage and contexts of the word or concept (e.g. in Figure 2).

Because of the different types of entries, there were several approaches used in selecting the examples. As per the recommendations by Landau (2001) and Pulverness & Tomlinson (2013), I used examples of real usage where possible. There are 692 examples in the AusDICT, an average of two examples per entry.

For the lexical items and phrases, examples were drawn from relevant media. As Landau (2001) points out, and the user focus

groups reiterated, examples of real speech are often too difficult for lower levels of language ability. As a result, in addition to real examples from written and spoken corpora, in some places I have included constructed scenarios illustrating the composition, as well as constructed examples adapted from real-life conversations, simplified examples from real speech and writing, and adapted phrases from social media.

Collecting and writing these examples were also one of the main challenges of this project and are discussed again below (see section 4.5.2).

4.4.4. Cross-references

An important element of this dictionary is the cross-referencing. I decided to do this with three different types of tags—"related values", "related words", and "related phrases". Each related term is provided as a hyperlink at the bottom of each entry. I used three different types of related terms in order to draw attention to the fact that some entries are definitions of specific terms ("related words" and "related phrases"), while others are descriptions of invisible culture—expectations, reactions, and other thought patterns ("related values"). They do not necessarily represent synonyms or antonyms, but rather words which are related to the current concept and enhance understanding, determined by either their relationships in previous research or in similarities emerging from the compositions themselves. In the ebook, these were three different colours to illustrate their differences; however, in the web version, the three types of relationships appear identical to the user (f. in Figure 2).

Of course, for all these related terms, it is impossible to be exhaustive, even within the AusDICT. All of language and culture is interrelated, so the relationships mentioned are determined by what is in the dictionary, what is relevant to the users, and what is necessary for understanding. Due to the several different pathways into the dictionary, the entries in each section were also cross-referenced to each other, despite their proximity and relationship being articulated by section headings. There are 1,500 related terms in the AusDICT, an average of 4.5 terms per entry.

4.5. Challenges of creating NSM-based cultural dictionaries

The main challenge of creating cultural dictionaries is that there are few extant cultural dictionaries with a focus on invisible culture to base features on and therefore every cultural dictionary project at this

stage is unique and faces the same challenges in uncovering and designing for user needs. This section will briefly discuss three of the main challenges which arose during the AusDICT project, and which are relevant for others designing cultural dictionary projects.

4.5.1. Challenge 1—Choosing headwords

Arguably the most recognizable part of a dictionary—the headword—determines what a user searches for in a dictionary. Throughout the AusDICT project, determining the most useful headwords for complex concepts and uncommonly expressed concepts was an important task. For the most part, headwords were determined based on the existing titles of explications and cultural scripts, and in consultation with the teachers during focus groups. In traditional dictionaries, there are many entries for each of the different senses of a word, under a single headword. In this dictionary however, there is only a single sense per headword, so the user must be able to determine immediately if the entry contains the information they are looking for. My solution to this challenge was to follow each headword for a lexical item with an example of usage (in NSM parlance, the "frame"). The simple example confirms the context for the user. Where different senses of a word are given, they appear as separate headwords.

For the entries describing cultural content, a more complex headword is needed. These headwords are mostly descriptive of a concept and use several words to express the idea. In general, I avoided using single words as headwords for cultural values so that the cultural values would not be mistaken for definitions of lexical items. In NSM research, titles of cultural scripts are useful indicators, but are not as useful for a layperson or dictionary user, as they can be relatively technical and defeat the purpose of using NSM and minimal languages as an approach. An example of this is the headword "valuing presumed social similarity and social equality". This is a challenge which I do not think I have found an ideal solution for, but have tried to address through the cross-referencing and navigation options, discussed below.

4.5.2. Challenge 2—Choosing examples

As discussed above, there were some practical challenges and adaptations made to selecting examples, but there was an additional challenge to including examples which required a more creative solution. The challenge was that cultural scripts that referred to values

or situations were difficult to provide examples for, as there were no specific linguistic features to include. For the AusDICT project I provided constructed examples for cultural values and attitudes, trying to keep scenarios short and interactional so that they provided enough context without oversimplifying the situation. Some of these examples are in minimal English, some are small role-play scripts, and some are both. This is not the ideal approach but was adopted based on feedback from the user testing which indicated that simple constructed scenarios and examples were useful for classroom contexts.

4.5.3. Challenge 3—Dictionary navigation

The final challenge I will discuss here is how users approach and navigate the dictionary to find entries. Most of the target users for the AusDICT are well-versed in using both paper and online dictionaries for which the first step in looking up an entry is knowing the word to search for. In the AusDICT, this approach works for the lexical items, but for values, attitudes, phrases, and norms, it is not always clear what the phrase or expression used as a headword is. My best solution to this was to continue the theme of cross-referencing and hyperlink elements to one another in as many ways as I could. In Figure 2 this is visible in the entry as f., g., and h., and in the three different menu options i., j., and k., as well as the search function n. In this way I have tried to make more visible what Tarp (2008) calls "invisible lemmas", or additional hidden headwords which improve searches. The search function also capitalises on this idea and a search term will additionally return terms related to the initial search so that a user can be swiftly guided to the most relevant content.

5. Final comments

While NSM-based cultural dictionaries are an emerging addition to lexicographical practice, there are plenty of arguments for why they are useful for both language learners and other users of a language. In particular, cultural information should not be as separate from definitions as it has been in the past and connecting the two more explicitly is a better representation of language. For users, having this knowledge allows users of the dictionary to become more competent and empathetic language users.

While scholars over many years, both within and outside of NSM, have advocated for including more cultural information in lexicographical works, it is the current capacity of digital and web-based lexicography which removes many of the barriers to cultural

dictionary-making. As with all emerging endeavours, creating NSM-based cultural dictionaries has challenges and problems which need to be solved, but only by pursuing more projects can we hope to overcome these challenges and solve these problems. It is my hope that the AusDICT can be used as a foundation for further advances in NSM-based cultural dictionaries.

References

Atkins, Sue B.T. 2008. Theoretical Lexicography and its Relation to Dictionary-making. In T. Fontenelle (Ed.), *Practical Lexicography,* 31–50. New York: Oxford University Press.

Atkins, Sue B.T. and Michael Rundell. 2008. *The Oxford guide to practical lexicography*. New York: Oxford University Press.

Atkins, Sue B.T. and Krista Varantola. 2008. Monitoring dictionary use. In Thierry Fontenelle (ed.), *Practical lexicography: A reader*, 337–75. New York: Oxford University Press.

Barrios Rodríguez, María Auxiliadora. 2013. Functional macrocategory and semiautomatic inheritance of semantic features: A methodology for defining nouns. In Valentina Apresjan, Boris Iomdin, and Ekaterina Ageeva (eds.), *Proceedings of the 6th International Conference on Meaning–Text Theory*, 11–20. Prague: Charles University.

Barrios Rodríguez, María Auxiliadora. 2020. Minimal and inverse definitions: A semi-experimental proposal for compiling a Spanish dictionary with semantic primes and molecules. In Lauren Sadow, Bert Peeters, and Kerry Mullan (eds.), *Studies in ethnopragmatics, cultural semantics, and intercultural communication: Vol. 3. Minimal English (and beyond)*, 191–212. Singapore: Springer. https://doi.org/10.1007/978-981-32-9979-5_10

Bullock, David. 2011. NSM + LDOCE: A non-circular dictionary of English. *International Journal of Lexicography* 24(2), 226–40. http://doi.org/10.1093/ijl/ecq035

Bullock, David. 2014a. *Learn these words first*. Retrieved 5 August 2018 from https://www.learnthesewordsfirst.com

Bullock, David. 2014b. What is a multi-layer dictionary? *Learn these words first*. Retrieved 9 January 2019 from http://learnthesewordsfirst.com/about/what-is-a-multi-layer-dictionary.html

Bullock, David. 2021. Using minimal English (minimal Spanish, etc.) for non-circular learners' dictionaries. In Cliff Goddard (ed.), *Minimal languages in action*, 111–37. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-64077-4_5

Cummings, Patrick J. and Hans-Georg Wolf. 2011. *A dictionary of Hong Kong English: Words from the Fragrant Harbor (Hong Kong English).* Aberdeen: Hong Kong University Press.

Fuertes-Olivera, Pedro A. 2010. *Lexicographica*. Series Maior: Specialised Dictionaries for Learners. Mouton De Gruyter. https://doi.org/10.1515/9783110231335.0.17

Goddard, Cliff. 2011. *Semantic analysis*, 2nd edn. New York: Oxford University Press.

Goddard, Cliff. 2017a. Minimal English: The science behind it. In Cliff Goddard (ed.), *Minimal English for a global world*, 29–70. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-319-62512-6_3

Goddard, Cliff. 2017b. Natural semantic metalanguage and lexicography. In Patrick Hanks and Gilles-Maurice de Schryver (eds.), *International handbook of modern lexis and lexicography,* vol. 20, 1–15. Berlin, Heidelberg: Springer Berlin Heidelberg. http://doi.org/10.1007/978-3-642-45369-4_14-1

Goddard, Cliff. 2018. A semantic menagerie: The conceptual semantics of ethnozoological categories. *Russian Journal of Linguistics* 22(3), 539–59. https://doi.org/10.22363/2312-9182-2018-22-3-539-559

Goddard, Cliff (ed.). 2021. *Minimal languages in action.* Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-64077-4

Haiman, John. 1980. Dictionaries and Encyclopedias. *Lingua* 50(4), 329–57. https://doi.org/10.1016/0024-3841(80)90089-3

Hirsch, E.D., Joseph H. Kett and James S. Trefil. 2002. *The new dictionary of cultural literacy*. Houghton Mifflin Harcourt.

Khalidi, Sara, Fiona McIlroy and Hannah Neumayer. 2012. *Cultural dictionary and directory*, 3rd edn. Migrant and Refugee Settlement Services of the ACT Inc. Retrieved 3 January 2019 from http://www.picacnsw.org.au/data/resources/The_Cultural_Dictionary.pdf (site discontinued, now available at https://uploads-ssl.webflow.com/602d05b9f629510c1d8b24ce/602fecdbd62c7756e566e3de_The_Cultural_Dictionary.pdf)

Landau, Sidney. 2001. *Dictionaries*, 2nd edn. Cambridge: Cambridge University Press.

Lew, Robert. 2015. Research into the use of online dictionaries. *International Journal of Lexicography* 28(2), 232–53. http://doi.org/10.1093/ijl/ecv010

Lew, Robert and Gilles-Maurice de Schryver. 2014. Dictionary users in the digital revolution. *International Journal of Lexicography* 27(4), 341–59. http://doi.org/10.1093/ijl/ecu011

Machin, Elita. 2021. Minimal English and Revitalisation Education: Assisting Linguists to Explain Grammar in Simple, Everyday Words. In Goddard, Cliff (ed.). *Minimal Languages in Action.* Cham: Palgrave Macmillan pp 83-107

Miller, Julia, Ardian Setiawan and Deny Kwary. n.d. *Australian cultural dictionary*. Retrieved 8 January 2019 from http://www.culturaldictionary.org/ (site discontinued)

Peeters, Bert. 2019. From cultural to pedagogical scripts. In Cliff Goddard (ed.), *Minimal languages in action.* Cham: Palgrave Macmillan.

Peeters, Bert. 2000a. Setting the scene. In Bert Peeters (ed.), *The lexicon–encyclopedia interface*, 1–52. Brill Academic Pub. https://doi.org/10.1163/9780585474465_002

Peeters, Bert (ed.). 2000b. *The lexicon–encyclopedia interface*. Elsevier. https://doi.org/10.1163/9780585474465

Peterson, Brooks. 2004. *Cultural intelligence*. Yarmouth: Intercultural Press.

Pulverness, Alan, & Tomlinson, Brian. 2013. Materials for Cultural Awareness. In Brian Tomlinson (ed.), *Developing materials for language teaching*, 443–460. Bloomsbury.

Sadow, Lauren. 2021. Standard Translatable English: A Minimal English for Teaching and Learning Invisible Culture in Language Classrooms. In Goddard, Cliff (ed.). *Minimal Languages in Action,* 139–169. Cham: Palgrave Macmillan.

Sadow, Lauren. 2020a. Principles and prototypes of a cultural dictionary of Australian English for learners. In Lauren Sadow, Bert Peeters, & Kerry Mullan (Eds.), *Studies in ethnopragmatics, cultural semantics, and intercultural communication: Vol. 3. Minimal English (and beyond),* 165–190. Singapore: Springer.

Sadow, Lauren. 2020b. Minimal English: Taking NSM 'out of the lab'. In Lauren Sadow, Bert Peeters, & Kerry Mullan (Eds.), *Studies in ethnopragmatics, cultural semantics, and intercultural communication: Vol. 3. Minimal English (and beyond),* 191–212. Singapore: Springer.

Sadow, Lauren. 2018. Can cultural scripts be used for teaching interactional norms? *Australian Review of Applied Linguistics, 41*(1), 91–116. DOI: 10.1075/aral.17030.sad

Sánchez, Aquilino. 2010. The Treatment of Cultural and / or Encyclopaedic Items in Specialised Dictionaries for Learners. In *Lexicographica. Series Maior : Specialised Dictionaries for Learners*, 107–131. Mouton De Gruyter.

Sharifian, Farzad. 2011. *Cultural conceptualisations and language*. John Benjamins Publishing. https://doi.org/10.1075/clscc.1

Sharifian, Farzad. 2014. Cultural linguistics. In Farzad Sharifian (ed.), *The Routledge handbook of language and culture*, 473–92. Routledge.

Silverstein, Michael. 2006. Old wine, new ethnographic lexicography. *Annual Review of Anthropology* 35(1), 481–96. http://doi.org/10.1146/annurev.anthro.35.081705.123327

Summers, Della. (ed.). 2006. *Longman dictionary of contemporary English*, 4th edn. Pearson Education.

Tarp, Sven. 2008. *Lexicography in the borderland between knowledge and non-knowledge.* Mouton De Gruyter. https://doi.org/10.1515/9783484970434

Wierzbicka, Anna. 1985. *Lexicography and conceptual Analysis*. Ann Arbor: Karoma.

Wierzbicka, Anna. 1987. *English speech act verbs*. Sydney: Academic Press.

Wierzbicka, Anna. 1992. Back to definitions. *Lexicographica* 8, 146–74.

Wierzbicka, Anna. 1996. *Semantics: Primes and universals*. Oxford: Oxford University Press.

Wierzbicka, Anna and Cliff Goddard. 2013. *Words and meanings*. Oxford: Oxford University Press.

Appendix

Cambridge Dictionary Online

Dogs: a common animal with four legs, especially kept by people as a pet or to hunt or guard things https://dictionary.cambridge.org/dictionary/english/dog?q=dogs (Retrieved 14 September 2022).

Dogs (Goddard 2018: 549–51)

a. creatures [m] of one kind
    they are animals [m], at the same time they are not like animals [m] of other kinds

b. they want to do many things with people, they can do many things as people want
    many of them live with people because people want this
    many of them are like this: when people say some words to them because these people want them to do something, they can know what these people want to say

c. some of them are big, someone can't pick up [m] one like this
    some of them are small, someone can pick up [m] one like this with two hands [m]
    some are very small, some are very big

d.   their bodies are like this:
—   one part of the head [m] sticks out [m], this part is below the eyes [m]
  the nose [m] is part of this part of the head [m], the mouth [m] is part of
   this part of the head [m]
—   their ears [m] are on two sides of the top [m] of the head [m]
—   they have sharp [m] teeth [m]
  they have long [m] tongues [m], often people can see parts of their
   tongues [m]
—   they have two legs [m] at the front [m] of the body
  they have two legs [m] at the back [m] of the body
—   some of them have long [m] legs [m], some don't have long [m] legs [m]
—   many of them have a long [m] tail [m], some don't have a long [m] tail [m]

e.   they can do some things with the mouth [m], when they do this,
  people can hear something of one kind because of it, someone far away
   can hear it
when people hear something of this kind, they can think about it like this:
  "it wants to say something like this: 'something is happening here now,
  I feel something now because of it, I want to do something now because
   of it'"

they can do some other things with the mouth [m], when they do this,
  people can hear something of another kind because of it
when people hear something of this other kind, they can think about it like this:
  "it wants to say something like this to someone: 'I feel something bad
   towards you,
  I want to do something bad to you with my teeth [m]'"

they can do some other things with the mouth [m]
when people hear something because of this, they can think about it like this:
  "it is feeling something very bad now"

f.   they can do things like this:
—   when they want to be somewhere else after a short time,
    they can move quickly [m]
—   they can do many things with their mouths [m]
  they can bite [m] other creatures [m], they can bite [m] people
  they can pick up [m] something with their mouth [m],
  they can hold [m] something with their mouth [m]
—   often when they want to eat [m] something, they want to eat [m] meat [m]
—   sometimes they kill [m] creatures [m] of other kinds

—   when they feel something good towards someone,
    they often do something with their tails [m]
  when they do this, the tail [m] moves many times in a short time

— when one of them is in a place, it can know many things
about this place because it can do something with the nose [m]
if someone was in this place not long before, it can know it

g. many people feel good things towards them
— many of them live in people's houses [m] because these people
want this,
many of them live near people's houses [m] because these people
want this
— often these people think about such animals [m] like this: "this is
someone"
they do some good things for this animal [m]
they often want it to know that they feel something good towards it
at the same time they want to know that it feels something good
towards them
— sometimes it is not like this, sometimes some people
do very bad things to animals [m] of this kind
— many animals [m] of this kind think about someone like this:
"this someone is not like any other someone"
when one of them thinks like this about someone, it wants to often
be with this someone, it wants to often do things with this someone

— sometimes animals [m] of this kind think about a place like this:
"this place is not like any other place"
when one of them thinks like this about a place, it wants other animals [m] not
to be in this place

— some animals [m] of this kind can do things of some kinds as people want
it can be good for people if these animals [m] do these things

h. when people think about animals [m] of this kind, they often think about them
like this:
"animals [m] of this kind are like no other animal [m]
when someone is with one of them, this someone can feel
something good
like someone can feel when this someone is with other people"
at other times they think about them like this:
"they are not like people
sometimes they do things with their parts of their bodies not like
people do
if people did such things, it would be very bad"