The World and Words of Mind

Bertram F. Malle
Department of Psychology
University of Oregon

What can be more fascinating, more central to psychology than the problem of capturing mental states with words—the magical relation between mind and language. Some of the brightest thinkers, perhaps even all significant philosophers of the 20th century, have at some point grappled with the question of how we are able to speak about the unobservable sphere of the mental, how we can actually refer to states of the mind. Sabini and Silver’s target article revolves around this question and thus draws attention to an essential problem of psychology that is all too often treated lightly or ignored. In fact, for a good 2 decades, Sabini and Silver have studied issues concerning the general problem of mind and language, and for that alone they should be commended. I admit that in many cases I don’t agree with their theories, their proposed solutions to the problems; but much more important is that they are working on problems that we as psychologists should care about.

In this target article, Sabini and Silver (this issue) take a step back from the specific analyses they have offered over the years about the nature of shame, guilt, pride, and related emotions. They propose instead a hypothesis that should hold for all emotion terms, presumably for all mental state terms. This hypothesis is that the mental state lexicon is richer than the reality of mental state experiences. More concretely, they suggest that there are “fewer emotional states than there are emotion words in the language” because there are several cases in which “the same mental state underlies the ascription of two different emotions.” This is a thesis about the mind–language relationship. The authors also draw the ontological conclusion that, because of the asymmetric mind-language relationship, “there are fewer unique mental states than one might have thought.”

There are certain theoretical commitments the authors make in formulating their thesis and in presenting the evidence to support it. These assumptions include, first, the conviction that “the emotion lexicon includes the names of mental states”. That is, emotion terms have a one-to-one correspondence with mental state names. This we can call the referential name commitment. However, the authors reject the belief that all emotion terms have a one-to-one correspondence with mental states. This is a key distinction, in that they treat M terms as names of mental states but do not think that each M term names exactly one experiential state. Their second important commitment is mental realism—the position that mental states actually exist in the mind or organism of the actor and that there are particular mechanisms and conditions that generate these states.

Sabini and Silver (this issue) then try to do two things in the target article. They say up front why they believe the mental lexicon might be richer than the mental world, and then they offer a variety of examples, arguments, and empirical data in support of their general thesis. Although, I believe the general thesis fails, the reasons why it fails are interesting in their...
General Argument: The Influence of Nontarget Information

Sabini and Silver (this issue) suggest that speakers decide how to describe experiences in response to three sources of information. There is the target person said to be having the emotion, the speaker doing the describing; and the audience for whom the description is given. The authors go on to emphasize that descriptions of a target’s emotions are heavily influenced by speaker and audience information. Such nontarget information includes, specifically, (a) what the speaker believes, wants, and feels; and (b) what the speaker thinks the audience believes, wants, and feels. The authors then suggest that emotion terms are so diverse not because of a corresponding differentiation of target experiences but because speakers also integrate considerable nontarget information into their emotion descriptions. So when my sister asks how my wife feels today, my choice of mental state terms in my answer will be influenced by my own goals, thoughts, and feelings as well as by my assumptions about my sister’s goals, thoughts, and feelings.

The first thing to note is that this claim appears to be at odds with the authors’ general thesis about the greater richness of mental state terms compared to mental states themselves. If the complex differentiation of speakers’ and audiences’ mental states (beliefs, desires, intentions, feelings, moods …) gives rise to the richness of mental or emotion descriptions, then the mental world can hardly be less complex than the world of mental descriptors. Counting might be difficult, but if every speaker’s choice among, say, three descriptors responds to the actual differentiation among, say, five speaker states and five audience states, then it would be rather difficult to argue that mental descriptors are more richly differentiated than mental states themselves.1 At the very least, it doesn’t follow from the above claim that mental descriptors are much more complex than the mental world.

Another, deeper problem is that the authors’ claim of “greater differentiation” sometimes applies to the meaning of emotion terms, sometimes to the nature of emotions themselves. At one point they argue that “distinctions among emotion terms” can depend on the speaker’s and audience’s mental states, not just the target’s, but in the same paragraph they argue that distinctions among emotions depend on those other factors. The latter is the far stronger thesis, claiming that the speaker’s and audience’s mental states can somehow alter the differentiation of emotions themselves. In the authors’ words, “what distinguishes one emotion from another might not be in the experience of the person having the emotional episode, but rather might lie in the experience of the person doing the describing or in the audience for the description.” But for realists, as Sabini and Silver (this issue) are committed to be, what distinguishes one emotion from another cannot lie in anybody’s description but in the characteristics of the emotions themselves. No further general argument supports this stronger version of the thesis, so I assume the authors stand only behind the weaker one.

The weaker thesis is that distinctions among emotion terms depend on the speaker’s (and the audience’s presumed) mental states. In one sense, this must be true. Speakers cannot always base their emotion descriptions solely on information about the target’s mind (often, e.g., they have no clear idea what is going on in that mind) and so emotion descriptions will be based on stereotypes or on guesses about what the audience wants to hear. Whether one says that June feels shame or guilt can of course depend on information wholly independent of what is genuinely going on in June’s mind. Nobody would debate this, so I again presume the authors are after something different, something a bit more controversial; but what?

To answer this question, I need to make a distinction between meaning and reference. Since Gottlob Frege (1892), this distinction has been illustrated with the difference between Morgenstern (morning star) and Abendstern (evening star). Both terms have the same referent—the planet Venus—but they differ in meaning. Morgenstern = the brightest star in the morning sky, and Abendstern = the brightest star in the evening sky. Now, speaker’s and audience’s mental states will of course influence the speaker’s decision to use Morgenstern or Abendstern. In a poem, for example, Abendstern might carry the bittersweet connotation of a beautiful life ending and Morgenstern, the connotation of a new life beginning. But none of these connotations will alter the fact that the speaker is referring to the same thing, the planet Venus.2 If we apply this model to emotion states, I might describe June’s mental state as shame or guilt depending what I think or what I want my audience to believe, but in each case I should refer to (or at least intend to refer to) the same mental state. So nothing necessarily follows from variations in emotion term meaning (due to speaker’s and audience’s mental states) for emotion term reference.

---

1Admittedly, the speaker’s descriptors are intended to describe the target’s mind, and my comparison here is with the speaker’s and the audience’s mind. But there is no basis for believing that a given target’s state of mind is simpler than a given speaker’s or audience’s state of mind. If speakers have to deal with the complex networks of mental states in self and audience, then surely they have to deal with the complex networks of mental states in the target as well.

2Unless, of course, the speaker falsely believes that there are two different stars—a complication treated in the philosophical literature (e.g., Kripke, 1972; Russell, 1905; Searle, 1958; Strawson, 1954)
but the latter is needed for Sabini and Silver’s thesis. In the extreme case, speakers of different languages will use different terms (with slightly different meanings) to intend to refer to the same mental state residing in June. But even counting such terms across different languages, we would not conclude that there is more complexity in the lexicon of mental state terms than in the world of mental states themselves.

Thus, Sabini and Silver’s (this issue) general argument about the significance of non target information on emotion descriptions appears to fail on two counts. For one thing, the argument presupposes a considerable complexity in the mental world of speakers and audiences that is hardly compatible with the intended conclusion that the linguistic world is more complex than the mental world. More important, the argument concerns the meaning and use of mental state terms whereas the authors’ hypothesis tries to make a statement about reference. The authors need to show that there can be a multitude of linguistic terms that pick out the same real emotion state, so they need an argument at the level of reference diversity, not just meaning diversity. The authors attempt to show such diversity in the second part of their article using specific pieces of evidence. Because there is an intimidating problem facing all such attempts, I introduce it first and then turn to the authors’ evidence.

**General Problem: The Identity of Mental States**

Sabini and Silver (this issue) observe that “the same experience may be differently described depending on the situation”. This must be true given the context-sensitivity of language and the fact that experiences don’t reliably determine a single (correct) description. But here is the problem. What do we mean by “the same experience”? We are speaking here of the real experiential state, independent of its description. From whose perspective do we assess the identity of this state, and using what criterion of identity? What guarantees that, speaking twice about an episode of falling in love, the speaker truly refers to the same episode each time?

One possibility is to adopt the speaker’s perspective and accept his or her identity criterion. Bruce had a special experience of meeting his wife. The subsequent episodic memories have changed over the years, as have the words each time he tells a story about the experience, but within each story he is clearly trying to refer to the same experience, to the best of his belief and memory. This subjective criterion is not incorrigible (as famously discussed by Wittgenstein, 1953). Bruce’s wife might correct him about details of the setting or remind him of what he said at the time, and Bruce may realize he is mixing up two experiences or engaging in a bit of embellishing. So another possibility is to simply adopt the best identity criterion that is socially available—which at times lies solely with the experiencer, at times incorporates witnesses of the event or previous audiences. It is Bruce’s task to use the best criterion available, but he cannot ensure that he is objectively referring to the “truly” same experience that occurred exactly once in the past. There is no such objective criterion, no fail-safe act of direct reference, neither from the 1st-person nor the 3rd-person perspective. In their discussion of specific evidence, however, Sabini and Silver appear to presuppose such an “objective” identity criterion, so let’s take a look at this evidence.

**Specific Pieces of Evidence**

**Marked Versus Unmarked Beliefs**

As a first, still somewhat general piece of evidence, the authors offer the explanation of Bill’s running away from the building as one illustration of the claim that “how one describes a situation involving emotion [actually, mental states] may depend on more than the experiences of the experiencing subject.” In this example, the speaker can decide to explain Bill’s behavior by saying “The building is on fire” or by saying “He thinks the building is on fire.” In each case (as my colleagues and I have shown; Malle et al., 2000) the speaker ascribes a belief to Bill, but in the first case the speaker takes the belief to be fact or at least expresses that he shares the belief. In situations like these, the speaker’s linguistic choice is indeed dependent on the speaker’s own goals and beliefs. But nothing follows for the authors’ general thesis about reference. In both cases the same mental state of “BELIEF [house is on fire]” is ascribed to Bill—once explicitly with a verb, once implicitly without one—but the speaker does not refer to one state with two distinct terms.

**Envy Versus Anger**

The first piece of empirical evidence the authors put forth in support of their thesis is an empirical vignette study in which the experimenters carefully constructed a contrast between how an actor himself would (probably) experience a situation and how spectators (i.e., participants in the vignette study) would describe the actor. The contrast is, specifically, between the actor’s feeling of righteous indignation and the spectators’ assessment that the actor feels envy (and participants do indeed recognize this contrast). From these findings the authors conclude that “the experiences of an envious person and of a rightly indignant person can be the same” and, a moment later, “righteous indignation and envy, we claim, are the same experiential state.” If true, this conclusion would be consistent with
the authors’ general thesis, as it says that two terms (righteous indignation and envy) refer to the same mental state. Note that the authors are not talking about meaning here—they don’t claim to have discovered that with righteous indignation and envy we have two previously unnoticed synonyms in our vocabulary. Rather, they claim that these two terms, despite their different meaning, can actually refer to or designate the same mental state.

Methodologically, I take this case to be more of an argument than an empirical test, because generalizing from a single, tailored vignette is always difficult. More problematic is the inference from a case of multiple names for one emotion state to the broad trend that the authors’ general thesis proclaims. What would follow, for example, from the case that there are millions of pebbles on the beach and I refer to one of them as a small rock whereas my conversation partner calls it a stone? People can of course name a particular thing (in a particular context) in different ways; this hardly proves that the number of words is generally larger than the number of objects named by those words.

But there is a much deeper problem. As we saw earlier, an identity criterion for two references to the same mental state is hard to come by; Sabini and Silver (this issue) certainly don’t offer us any. Their study on anger and envy merely illustrates a case in which the actor and the spectator differ in their descriptions of a vaguely bounded emotion-action complex. And not only do we lack solid evidence that each party is referring to just one identifiable “thing”; there is good reason to believe that the parties are actually referring to distinct aspects of that complex. By stipulation, the actor describes what he is consciously feeling at the moment; what the spectators describe is the whole emotion-cum-action with which they are presented (reflecting a rather general actor-observer asymmetry in attention and reference to experiences vs. actions; Malle & Knobe, 1997b; Malle & Pearce, 2001). No spectator could censure the actor by saying: “No, no, you feel envious!” At best, they might say “You are envious,” and that is not a description of a conscious experiential state but of a complex of causal histories, behaviors, and perhaps unconscious emotions. Consider this case: If spectator A said “He is envious” and the actor said (honestly) “I am not envious,” we would not conclude that envious and not envious are two terms for the same state. We would conclude that they disagree, or that one is right and the other is wrong, or that they are talking about two different things.

**Shame and Embarrassment**

The second piece of evidence is a vignette study that leads the authors to conclude that “shame and embarrassment are different interpretations of the same raw feeling”. In their analysis, Sabini and Silver (this issue) show us subtle aspects of the two emotions. But can they support their general thesis of more mental terms than mental states? Let’s take a closer look. In the vignette, a protagonist is being associated with a piece of pornography under two conditions: In one, he actually owns the piece, in the other, his office mate owns it. The authors’ thesis requires that these two protagonists are in the same “raw feeling state” or “raw physiology state” and the two terms shame and embarrassment actually refer to that identical internal state. We are thrown back into the problem of identity. The protagonist’s psychological world in each case is an interwoven complex of knowing, realizing, fearing, hoping, sensing, intending, and much more. Where exactly are the boundaries of the postulated “raw state”? Are the hot flashes that the “shameful” person feels and the embarrassed person lacks part of that raw feeling? Sabini and Silver treat the mental world as segmentable and its segments as uniquely identifiable, but this is a treatment borrowed from physical objects (for which intersubjective segmenting and identification are principally possible) and is far less suitable for psychological states.

**Regret**

The final piece of evidence concerns regret. Sabini and Silver describe Lord Jim’s case in which the term regret is an acceptable descriptor of the actor’s emotional state only if shame is an acceptable descriptor as well. In addition, rating data from Sabini and Perez (1996) suggested that regret is often associated with a variety of negative emotions. Whether or not Sabini and Silver (this issue) are correct in their further conclusion that regret is never used by itself (and therefore does not refer to a unique mental state), the postulate of combined emotion labels does not support the authors’ general thesis that there are more mental state terms than identifiable mental states; it actually undermines it. Let’s count. If there are, say, 50 distinct emotion labels that can be combined with each other, then their combination will be able to describe a much larger number of different emotion states than there are labels (e.g., 1225 states labeled by pairwise combinations of terms). People combine and recombine their finite number of mental state terms precisely because they seek to track the complexity and multifacetedness of mental states—a complexity that goes far beyond the available single-state terms that language provides.

**Conclusion**

Emotions are complex, the experiential and situational contexts in which they occur are varied, and we won’t find neat correspondence relations between labels and states. If that were the authors’ thesis, I would
mind–language relations, we need to return to some-
other minds. If we want to build a solid model of these
the human mind to represent and communicate about
the complexity of the mental world. Specific contexts pro-
vide criteria that allow people—actors, interaction
partners, or spectators—to assess the appropriateness
of the mental state terms used in this context. Those
criteria include introspection, memory, observation,
joint attention, logic, negotiation, and many more. That
is as far as “reference” relations between mind and lan-
guage go. There is no reliable way of carving up mental
complexes and singling out those elements that are
generally referred to with specific language terms.
Across contexts, mental state terms therefore have dif-
ferent relationships with mental states, just as in some
of the cases Sabini and Silver describe. But also none
of these specific cases can stand for the whole range of
meanings, uses, and references that mental state terms
accomplish.

Mental words are not names of things; they are so-
plicated attempts to capture the complexity of minds in a complex web of language. This language
does not consist of general laws of reference but of cre-
ative recombinations of a finite set of terms that allow
the human mind to represent and communicate about
other minds. If we want to build a solid model of these
mind–language relations, we need to return to some-
thing long out of fashion: A descriptive psychology
that catalogues the whole range of mental state terms
and the contexts in which they are used, including their
causal histories, connotations, and social functions.

Coda: The Study of Folk Concepts

What would a serious descriptive psychology look
like? The classic approach is one of careful linguistic
and conceptual analysis, primarily in the style of ordi-
ary language philosophy (e.g., Austin, 1961; Ryle,
1949; Wittgenstein, 1953; see also Heider, 1958). To a
large extent, this is the methodology Sabini and Silver
(this issue) are using, with the addition of tailored vi-
nette studies. And even though one has to admire the
authors’ sensitivity to subtleties of usage and context,
the complexities of mind-language relations require a
more systematic and less subjective approach (Malle,
The descriptive study of folk concepts—and shame,
guilt, pride, regret are clearly words that stand for folk
concepts—must employ the entire armada of psycho-
logical methodologies we have available: interviews,
surveys, experiments, video observation, archival
study, reaction time studies, and so on. But for starters
we also need a commitment to systematic empirical
data collection. The most sensitive intuitions about
mental state concepts can be debated and are often in-
correct; only empirical studies can tell us in which con-
texts people use which words of mind, with what crite-
ria of reference, and for what social purposes. This is
the reality psychologists must capture; and it is to
Sabini and Silver’s credit that they have reminded us of
this task.

Note

Bertram F. Malle, Department of Psychology, 1227
University of Oregon, Eugene, OR 97403–1227.
E-mail: bfmalle@darkwing.uoregon.edu

References

Austin, J. L. (1962). Philosophical Papers. CITY,
STATE/COUNTRY: Oxford.
Frege, G. (1892). Über Sinn und Bedeutung [On sense and refer-
ence]. Zeitschrift für Philosophie und philosophische Kritik,
100, 25–50.
Heider, F. (1958). The psychology of interpersonal relations. New
York: Wiley.
Davidson, Semantics of natural language (pp. XXX–XXX).
CITY, STATE/COUNTRY: Dordrecht.
Smedslund’s “The structure of psychological common sense.”
Contemporary Psychology, 43, 671–672.
Emotions and the Lexicon

James A. Russell
Department of Psychology
Boston College

How many kinds of emotion are there? And what are they? It seems natural to ask such questions, and to answer them with a list: anger, fear, shame, envy, embarrassment, and so on. Unfortunately, use of these English words as a scientific list of emotions faces a problem: Different languages categorize emotions differently. Although there are large similarities, especially among the Indo-European languages, there are also important differences, especially when non-Indo-European languages are considered (Russell, 1991; Wierzbicka, 1999). A speaker of Ifaluk (see Lutz, 1988) would come up with a different list.

Sabini and Silver’s (this issue) argument will fall on receptive ears. Many emotion researchers are already skeptical that language can be relied on to reveal much about emotion. In this regard, language is no different from any other symptom of emotion. No single index of emotion corresponds exactly to the emotion itself. George’s smirking facial expression depends not only on George’s emotional state, but also on his audience and context. (This claim is consistent with accounts of facial expression as different as Ekman’s, 1972, acknowledgment of display rules; Fridlund’s, 1994, audience effects; and my contextual-dimensional account, Russell, 1997.) Similarly, George’s peripheral physiological changes depend not only on his emotional state but on his intended action. And so on, for each single observable index of emotion, including self-perception of emotion, vocal expression, and instrumental behavior.

The problem that Sabini and Silver’s (this issue) raise about language appears familiar and appears to have a separate issue from one’s emotional state. Obviously, not all emotional reactions and the beliefs on which they are based are justified. Sabini and Silver (this issue) analyze the distinction between anger and envy and that between shame and embarrassment and, in each case, considerations beyond emotional state are involved. Their analysis shows that different lexical entries—words such as anger, fear, etc.—fail to correspond, one-to-one, to different emotions.

Sabini and Silver’s (this issue) raise another problem—based on the pragmatics of language use—in analyzing language to answer such questions. When someone sincerely attributes an emotion to another, that is, when someone honestly says “George is X,” where X is angry, afraid, ashamed, or any other emotion-descriptive word, the word chosen, X, varies with at least three factors: George’s emotional state, the speaker’s assumptions about George and his situation, and the speaker’s intended audience. Furthermore, and this is the important point, the last two factors can be lexicalized. What follows is that lexical differences can correspond, not to emotional differences, but also to differences in the speaker’s assumptions and intended audience. No lexicon, in English or any other language, maps directly onto emotion.

To illustrate: I see George as smug, whereas George sees himself as simply pleased. Self-attributes of smugness are rare, because smugness presupposes that the speaker finds George more pleased, more self-satisfied, than is warranted. Song is a word in Ifaluk that can be glossed as “justifiably angry” (Lutz, 1988)—close to but perhaps milder than ‘righteous indignation’ in English. Attributes of smugness or song involve considerations of justification, but justification is a