Disentangling the Affective Lexicon

Andrew Ortony and Gerald L. Clore

University of Illinois
at
Urbana-Champaign

When social psychologists and personality theorists investigate traits and emotions they frequently rely on lists of words that denote, or are thought to denote, traits, or emotions, or feelings. A classic problem in such work is that there is no unambiguous way of specifying which words refer to emotions, which refer to traits, and which refer to other behaviors, and non-emotional states. Investigators have generally relied on their intuitions in these matters, and by and large agreement has not been very high.

This problem is not so severe in lists of personality-trait words, largely because many modern empirical studies involving trait descriptors choose their terms from one of several "standardized" lists. For example, social psychologists frequently draw from the list of 555 words compiled by Anderson (1968). This list was developed by selecting feasible candidates from the 18,000 words appearing in Allport and Odbert's (1936) classic monograph. From the resulting reduced list of 2200 words, Anderson removed extreme words (e.g. majestic), words designating temporary states (e.g. aghast), words having to do with physical characteristics (e.g. hairy), strongly sex-linked words (e.g. alluring), and other words considered unsuitable as ingredients in impression formation (e.g. fond). Finally, words found to be unfamiliar to college students were eliminated. Although Anderson's list was also determined primarily on the basis of intuition, and although it does contain some ambiguous words (for example, happy, while certainly designating a trait, also can designate an emotion), it nevertheless has sufficient face validity to have gained wide acceptance.

Those who would study the emotions are less fortunate. Whether one seeks to map out the cognitive basis of the emotions, as we do, or whether one is investigating the effects of emotions on behavior, or of behavior on emotions, an indiscriminate use of language can be dangerously misleading in both theory construction and in the conduct of research. Many of the words in lists used in studying emotions either do not designate the kinds of states they are intended to, or they are ambiguous between different kinds of states. The indiscriminate use of such lists in theoretical and empirical research poses a serious methodological problem. For example, Russell (1980) scaled 28 "emotion-denoting adjectives". He found "sleepiness" to be an important dimension of such words. Although he lists the words used in his studies, he provides no justification for their inclusion and he describes no method for their selection. He included words like bored, tired, sleepy, drowsy, tranquil, and relaxed; we do not think that such terms denote emotions at all. If one includes among one's stimuli, words that have a high loading on sleepiness, sleepiness will turn out to be a factor. Until the inclusion of such words in the stimulus set can be justified, generalizations about the structure of the emotions have to be regarded as suspect.

The specification of necessary and sufficient criteria for emotions is a notoriously difficult if not an impossible goal. But because the employment of linguistic stimuli is an important avenue into the study of the emotions, some alternative method is needed for identifying emotion words. What we propose in this paper can be viewed as the linguistic groundwork required for language-based studies of the emotions and other affect-related terms.

The problem that we are dealing with is by no means unique to this domain. There are many areas where it is difficult or impossible to be entirely explicit about the criteria for class membership, but where psychologists have relied on rating scales and the intuitions of judges to achieve reliable and valid classifications. The work of Rosch and her colleagues (e.g. Rosch, 1978; Rosch & Mervis, 1975) on the categorization of concrete objects is an obvious example.

In the present context, one possibility would be to present subjects with candidate emotion words and ask how good good they are as examples of emotions. This we plan to do. However, as the sole strategy, this approach has drawbacks related to not knowing what criteria subjects employ in their judgments, and consequently it raises troublesome problems about reliability. Thus, as a first step, we chose to employ a number of explicit tests that we hope offer greater reliability and that provide potentially useful additional information about the structure of the affective lexicon. These tests constitute a set of heuristics for isolating genuine emotion words (and other kinds of words) from a list of putative emotion words. They take the form of a group of sentence frames into which a candidate word is inserted. The tests are "passed" or "failed" by a particular word depending on the extent to which groups of judges consider the resulting sentences to be meaningful rather than anomalous.

Finally, it should be emphasized that we think of these tests as a set of heuristics or "rules of thumb" rather than as an algorithm. Nevertheless, we think that they do a tolerable job of disentangling the affective lexicon--certainly a better job than blind intuition, or than no criteria at all.

While our primary goal is to isolate the genuine emotion words from a pool of candidate emotion words, we also consider it interesting to attempt to classify the major kinds of words appearing in the pool. The pool comprises the union of several lists claiming to be lists of emotions and/or of feelings. In constructing it we drew primarily from lists used in various psychological treatments of emotion (Bush, 1972; Dahl & Stengel, 1978; Davitz, 1969; Russell, 1980). The final pool consisted of about 500 words, the largest contribution coming from Dahl and Stengel's extensive list. A sizeable number also came from Bush (1972), who had reduced a prior list of 2,186 adjectives from Allport and Odbert (1936). From these Bush selected the 263 words that raters agreed were more relevant to emotions ("what a person feels") than to personality ("what a person is like") or to behavior ("what a person does"). Also included was Davitz's list of words from Roget's Thesaurus, and other smaller lists.

We found as we examined these lists that while we could not give a satisfactory definition of an emotion, we could readily eliminate many of the candidates as words that did not refer to emotions. For example, in the lists of words designating emotions and feelings used by Dahl and Stengel (1978) or Bush (1972), there are numerous "intruders" such as tired, hungry, breathless, and revived-- words which seem to designate body states, and words like confused, baffled, and sure, which seem to represent

non-affective cognitive states. Still other entries like abandoned, abused, and appreciated represent the acts or beliefs of others relevant to the self; they could certainly cause emotions but do not themselves denote emotions.

The linguistic tests that we propose are attempts to classify such "intruders" in a reasonably systematic way while also separating out emotion words. The first distinction we make is between words that designate traits or emotions and words that do not. Words that do designate traits or emotions are of three kinds: (a) "pure" trait words, which refer only to traits (and not to emotions), (e.g. studious, ambitious, mean), (b) "pure" emotion words, which refer only to emotions (and not to traits) (e.g. jubilant, distressed, embarrassed), and (c) polysemous words that can be used to refer to both emotions and traits (e.g. cheerful, happy, proud). For brevity we shall refer to such words as "emotion-trait hybrids". Although less central to our concerns, still of interest are the three kinds of words already mentioned that constitute the other half of the pool. These we call "body-state" words, "cognitive-state" words, and "other-action" words. The tests that we discuss in this paper are all designed to deal with adjectives or participial forms. Rephrasing of the tests is required to handle noun and verb forms.

Words denoting emotions or traits

The first test that we propose is actually a pair of sentence frames. One, Frame A, deals with negatively valenced words, and one, Frame B, deals with positively valenced ones. These frames can be thought of as linguistic filters. Their logic is to contrast candidate words with something explicitly emotional so that words that do not denote emotions will produce meaningful (as opposed to anomalous) sentences. The test separates the entire pool into two halves: (a) an item that fails the test (i.e. produces an anomalous sentence) is most probably a word that denotes a trait or an emotion, and, (b) an item that passes the test (i.e. produces an acceptable sentence) is probably a body-state word, an other-action word, or a cognitive-state word. Thus, the test is intended to allow as sensible completions only words like puzzled and certain (cognitive-state words), breathless and refreshed (body-state words), and abandoned and appreciated (other-action words).

Test 1.
Frame A: Although at that moment Mary was xxxxx, she was emotionally content
Frame B: Although at that moment Mary was xxxxx, she was not emotionally content

The word although anticipates a contrast, and in the contexts of these frames, it is a contrast of valence. However, the presence of the phrase emotionally content, constrains the contrast to non-emotional terms.

Accordingly, emotion words will fail the test, but body-state words, cognitive-state words, and other-action words all pass it. For example, words like breathless, puzzled, and abandoned pass the test because they fit the sentence frame for negative words (Frame A), and words like refreshed, certain, and appreciated pass because they fit the frame for positive words (Frame B). Traits are prevented from fitting into the sentence frames by incorporating in the frames a reference to a particular moment so that a quality that is enduring will give rise to an anomalous sentence. Thus, trait words as well as emotion words fail the test (e.g. honest, unkind, jubilant, and distressed).

Since our primary goal is to separate trait descriptors from emotion words, we shall deal first with that part of the initial pool that fails Test 1. Recall, first, that terms like proud, sad, and happy are sometimes used as trait descriptors and sometimes as emotion words. Thus, the half of the pool containing traits and emotions actually contains words of three kinds--the "pure" emotion words that unambiguously designate emotions (e.g. embarrassed, disgusted, jubilant), the "pure" trait words that unambiguously designate traits (e.g. thrifty, intelligent, studious, dishonest), and the "emotion-trait hybrid" words that have two senses, one referring to an emotion and one to a trait.

The test that we now describe is designed to separate pure trait terms and emotion-trait hybrids from pure emotion terms. Because the context provided by the sentence frame resists temporary states in favor of perservering qualities, it allows as sensible completions only traits and hybrid words with a trait as one meaning.

Test 2: John was well-known as a(n) xxxxx person

The result of applying this test is to separate examples like the following:

(PASS)	(FAIL)
pure traits and hybrids	pure emotions
anxious	disgusted
happy	distressed
proud	embarrassed
materialistic	jubilant
superstitious	love-sick

In order to separate the hybrids from the pure traits, another test, Test 3, is needed. This test may be applied to the same set of words as Test 2. The hybrids can then be isolated by taking the intersection of words passing Test 2 and of those passing Test 3. This is because Test 2 detects words that have trait readings, while Test 3 detects that subset of them that also have emotion readings (see Fig. 1).

The rationale behind Test 3 is that emotions can be experienced to varying degrees, and can be experienced in the absence of an interpersonal exchange. Thus, reflecting on a situation can give rise to an emotion but not to a trait, although, if a term is ambiguous as between a trait and an emotion, it will fit the test because of its emotion sense.

Test 3: As he reflected on what had happened, John was quite xxxxx

The result of applying this test is to separate examples like the following:

(PASS)	(FAIL)
pure emotions and hybrids	pure traits
cheerful	ambitious
distressed	intelligent
disgusted	knowledgeable
ecstatic	mean
frightened	sensitive
proud	thrifty

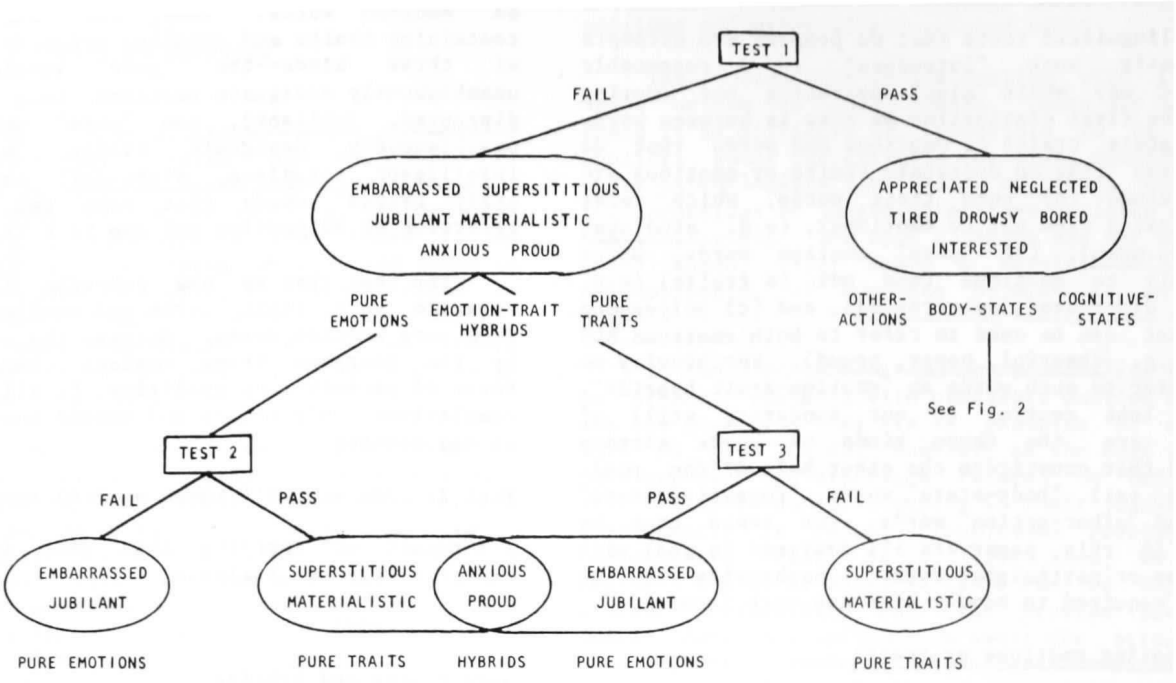

FIGURE 1.

Our primary goal has now been achieved. We think we have proposed a reasonably methodical procedure for isolating emotion words from a pool that contains some words that do not denote emotions. Furthermore, we have separated two kinds of emotion words, those that seem to denote emotions exclusively, and those (hybrids) that also denote traits. We consider this to be a potentially important distinction. The results of multidimensional scaling studies, for example, in which subjects make similarity judgments can be muddied by the unwitting inclusion of a subset of ambiguous stimuli (i.e. hybrids).

Words not denoting emotions or traits

There may be occasions on which one might want to compare emotion words to some other kinds of words, say, cognitive-state words, or other-action words. Although of secondary interest to us, the same kind of procedures can be used to separate the three kinds of words appearing in the other half of the initial pool, namely the half comprising words that passed Test 1 (see Fig. 2).

The first kind of words that we attempt to isolate are those that do not represent an internal state of a person at all. These we call "other-action" words because they characterize the actions (or the attitudes) of others that are relevant to (although not necessarily directed towards) the self. Perhaps because they are so strongly associated with emotional responses many other-action words have found their way into lists of emotions and traits. For example, the word abandoned appears in the lists of Dahl and Stengel (1978) and of Bush (1973). It also appears in the Personal Traits column of Allport and Odbert's (1936) list. However, in modern English abandoned designates neither a trait nor an emotion. One cannot be disposed to behave "abandonedly", (although we do speak of behaving "with gay abandon" meaning wrecklessly), and one does not experience "being abandoned" as a separate emotion. Rather, abandoned represents the actions of some other vis a vis the self. Its special quality, its emotional loading, presumably comes from the fact that the knowledge that one has been abandoned typically gives rise to (negatively toned) emotions. It is, however, as much of an error to assume that "abandoned" is an emotion or a trait as it would be to suppose that "kicked in the groin" was.

We shall assume that unlike emotion words, the most salient characteristic of other-action words is that others can engage in those actions without the person to whom they are relevant necessarily being aware of them. Since one can be abandoned and not know it, abandoned cannot represent an emotion or any other kind of internal state; it is an other-action word. Notice that it does not follow from this that awareness entails an emotional state. Normally, awareness is a necessary but not sufficient condition for an emotion. Thus, Test 4 is designed to identify other-action words. The logic of the test is (a) to deny awareness by using the expression "totally unaware", and (b) to take advantage of the fact that "other-action words" require actions by others that might influence the self by explicitly making the agent of the action an other.

Test 4: John was totally unaware that he
had been xxxxx by the woman

For this sentence frame reasonable completions are restricted to words that denote the actions (physical or mental) of others. As with the earlier tests, some words will fail to fit simply because they are of the wrong syntactic type, but, as always, the more interesting cases are those for which the resulting sentence is not syntactically ill-formed but

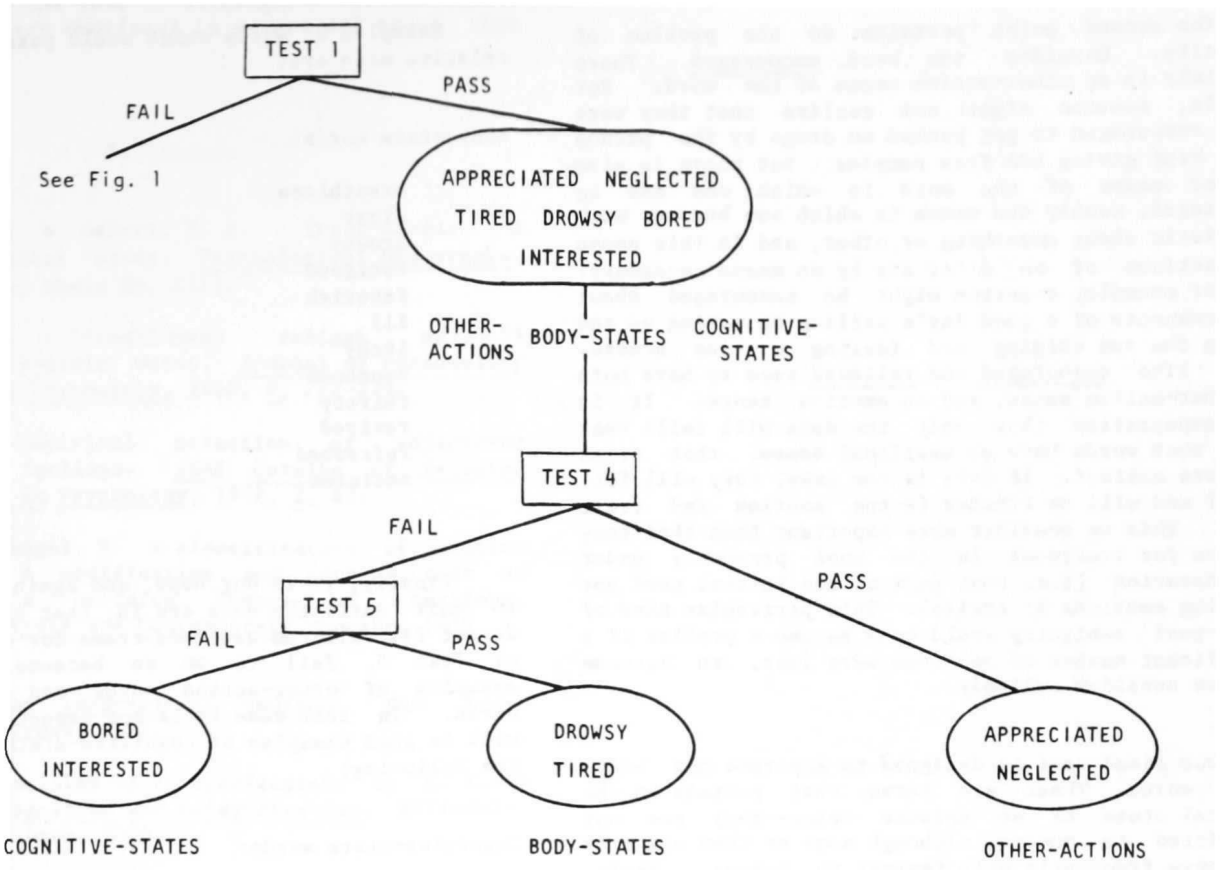

FIGURE 2.

rather is semantically anomalous. For example, the word puzzled does not fit very well because it is odd to suppose that John could be puzzled and not realize it. We think that puzzled fits better into the category that we call "cognitive-state" words. A more difficult example is revived. Because revived suggests the possibility of prior unconsciousness it seems better able to fit into the sentence frame, yet we like to think that revived is a body-state word. If this is so, then subjects should judge it to fit better in the body-state frame (see Test 5, below), even though it might do tolerably well in the other-action frame. Since subjects are asked to judge how well a target word fits in a frame, it would be sufficient for our purposes to discover that words like abandoned and ignored fit better than words like revived. It is not necessary that words not in the category upon which we are focussing give rise to seriously anomalous sentences. Our expectation is only that the most reasonable completions are produced by other-action words. Internal state words make poor completions. Thus, we can separate other-action words using this test. When subjects in our experiments are instructed to make meaningfulness judgments they are warned to ignore one particular interpretation of the sentence-frame in Test 4 that would confound the results by rendering spurious "meaningful" judgments. Subjects are told that the focus of the sentence should be on John's lack of awareness, not on the identity of the agent responsible for the action. Thus, they are instructed to ignore the interpretation in which John might have wrongly attributed his being ignored, appreciated, revived etc. to someone other than the woman. The following are examples of words that we think pass Test 4 most easily.

Other-action words:

abused
abandoned
appreciated
defeated
disgraced
ignored
neglected
slighted
welcome

It is worth pointing out a couple of things at this juncture. The first concerns the difference between "being" and "feeling" something. The inclination to treat other-action words as internal state words is much greater when they occur with "feel" than with "is". The reason is that "feel" can be, and often is interpreted to mean "feel as one would if (one realized that) one was xxxxx". Thus, that John was ignored entails nothing about what John felt. It merely asserts that somebody ignored John. Whether or not John responded to this other-action emotionally will depend on all kinds of factors (e.g. Was he aware of the fact? Did he expect anything else? Did he care? etc.) In other words, the inference to an emotional response is a pragmatic one, not a logical one. Yet, if one says that "John felt ignored", we have much more license to infer that John was in an (emotional) internal state. We infer that John responded emotionally. There is no doubt that feeling ignored is a unique kind of (negative) feeling--so too is the feeling of being pricked by a needle. But this fact is not sufficient for it to

count as an emotion. One cannot confuse causes with their highly correlated effects.

The second point pertains to the problem of ambiguity. Consider the word encouraged. There certainly is an other-action sense of the word. For example, someone might not realize that they were being encouraged to get hooked on drugs by the person that kept giving him free samples. But there is also another sense of the word in which one can be encouraged, namely the sense in which one becomes more optimistic about something or other, and in this sense the actions of an other are by no means necessary. So, for example, a person might be encouraged about the prospects of a good day's sailing on waking up and seeing the sun shining and feeling a fine breeze. Words like encouraged and relieved seem to have both an other-action sense, and an emotion sense. It is our expectation (but only the data will tell) that where such words have an emotional sense, that sense is more salient. If this is the case, they will fail Test 1 and will be treated in the emotion and trait pool. This we consider more important than that they survive for treatment in the pool presently under consideration (i.e. that part of the initial pool not denoting emotions or traits). This particular kind of cross-pool ambiguity would only become a problem if a significant number of emotions were lost, an outcome that we consider unlikely.

Our final test is designed to separate out body-state words. These are terms that pertain to the physical state of an animate being--they are not restricted to humans, although some of them might be used more frequently with respect to humans. Again, these words can be valenced, and are often, but by no means necessarily, associated with emotional responses. Their appearance in various lists of emotion words (e.g., again, those of Dahl & Stengel, and of Bush, and of Russell) is probably due to the fact that they appear in (Column II of) the Allport and Odbert (1936) list. This category is loosely characterized by these authors as "terms designating mood, emotional activity, or causal and temporary forms of conduct" (p.vii). In it appear words like thirsty and breathless which in our opinion do not fit even this loose characterization. What is it to be in a thirsty mood? Is being thirsty an emotional activity, or a form of conduct? We suspect that these terms appear in Column II not because they belong there, but because they are less incongruous there than in one of the other three categories used by Allport and Odbert.

Test 5 attempts to separate out these terms by using a sentence frame that focuses on body states (as opposed to other kinds of sensations, or perceptions), and that minimizes the cognitive content by predicating them of a newborn infant:

Test 5: The pediatrician explained that one of the physical characteristics of a newborn infant was to be xxxxx.

It seems to us that this test only allows as good completions terms that designate body feelings. It seems to us to more readily allow completions with words that do not entail awareness and that do not suggest cognitive activity (emotional or otherwise). Thus it would be odd to complete this frame with an other-action word like ignored, and it would be odd with words like certain. The oddness arises both from attempting to predicate higher level cognitive functions involving social awareness and metacognition to newborn infants, and from the fact that these predicates do not refer to physical characteristics.

Examples of words which would pass this test with relative ease are:

Body-state words:

breathless
dizzy
drowsy
fatigued
feverish
ill
itchy
nauseous
thirsty
revived
refreshed
satiated

Finally, it is our hope, and again we shall have to wait until the data are in, that those words that do not fit well the sentence frame for either Test 4 or Test 5, fail to do so because they are poor examples of other-action words and of body-state words. In that case it is our expectation that they will be good examples of cognitive-state words such as the following:

Cognitive-state words:

bored
disbelieving
distracted
doubtful
overworked
puzzled
uncertain
uninspired
amused
aware
certain
impressed
interested
sure
vindicated

Conclusion

The question of the psychological validity of the various categories that we have proposed is obviously an important issue. We find these categories to be intuitively reasonable and we believe that they do represent psychologically important distinctions. However, ultimately we would like to know that these distinctions correlate with behavioral differences. For example, in a pilot study conducted by Lord and Ortony memory for emotion words was found to be very much superior to memory for cognitive-state words. These are the kind of data that one needs to demonstrate the psychological validity of the distinctions we have proposed.

Finally, we should point out that we are more wedded to the general principles that we have proposed than we are to the specific tests. Indeed, some of the tests we find rather inelegant. It remains to be seen how effective these tests are, and we are convinced that there is room for considerable improvement. However, some procedure along the lines of the one we have suggested seems essential if one is to avoid the kinds of problems in the analysis of

emotions that we identified at the outset. We hope that our discussion will alert those who are studying the emotions to the need to distinguish between states that genuinely are emotional in nature and those that are not.

References

Allport, G. W., & Odbert, H. S. Trait-names: A psycho-lexical study. Psychological Monographs, 1936, 47(1, Whole No. 211).

Anderson, N. H. Likeableness ratings of 55 personality-trait words. Journal of Personality and Social Psychology, 1968, 9, 272-279.

Bush, L. E. Empirical selection of adjectives denoting feelings. JSAS Catalog of Selected Documents in Psychology, 1972, 2, 67.

Dahl, H., & Stengel, B. A classification of emotion words: A modification and partial test of de Rivera's Decision Theory of Emotions. Psychoanalysis and Contemporary Thought, 1978, 1, 269-312.

Davitz, J. The language of emotion. New York: Academic Press, 1969.

Rosch, E. Principles of categorization. In E. Rosch (Ed.), Cognition and categorization. Hillsdale, N.J.: Erlbaum, 1978.

Rosch, E., & Mervis, C. B. Family resemblances: Studies in the internal structure of categories. Cognitive Psychology, 1975, 7, 573-605.

Russell, J. A. A circumplex model of affect. Journal of Personality and Social Psychology, 1980, 39, 1161-1178.

Acknowledgements

We would like to thank Charles Lord and George Miller for helpful comments on earlier drafts of the paper. The research reported herein was supported in part by the National Institute of Education under Contract No. HEW-NIE-C-400-76-0116, and in part by a Spencer Fellowship awarded to the first author by the National Academy of Education.

Affect amd the Perception of Risk

Amos Tversky and Eric J. Johnson
Stanford University

As cognitive scientists turn their attention to emotion, they face the task of integrating affect into models of cognition. The perception of risks seems an ideal area to examine the relationship between cognitive and affective processes. When we witness an accident, or read a newspaper story about a natural disaster, we do more than simply revise our subjective probabilities. We are often quite disturbed and shaken by such events. Our encounters with risk are inevitably connected with feelings, including those of surprise, dismay, and worry.

Previous work in risk perception has concentrated on the cognitive domain. Lichtenstein et al. (1978), for example have asked people to estimate the frequency of death due to various causes. They argue that the availability of instances in memory helps determine these perceived frequencies. Thus, homicide is seen as much more common than suicide, although actually the reverse is true. Causes of death which are spectacular and the subject of media coverage appear to be overestimated while more mundane causes are underestimated.

We conducted three studies using an experimental paradigm similar to the one used by Lichtenstein et al. Before they made their estimates, however, subjects read a newspaper-like account of the death of a single individual under the guise of a newspaper reporting study. These stories, although quite graphic, were relatively devoid of information. They were, however, effective in changing mood, causing readers to report they felt much more depressed than a control group which had not read the stories. Later, in an apparently unrelated questionnaire, these subjects were asked to estimate the frequency of death due to various causes. The causes of death ranged from those closely related to the topic of the story, such as stomach and lung cancer for a story about leukemia, to unrelated causes such as tornados and airplane accidents.

The potential impact of these stories, and their accompanying changes in mood, represent a continuum. At one end of the continuum, we might expect the story to have no effect on the estimates. This is the normatively justified response, since the stories contained no information about the frequency of the death in the population. In contrast, the reader of the story might generalize from the instance in the newspaper-like story and increase their estimate of the frequency of that cause of death. We will term this a local generalization.

The impact of the story might also generalize to other, related risks. A story about a leukemia victim might also raise our subjective probability of related diseases such as lung and stomach cancer, but not unrelated risks such as airplane accidents. This gradient generalization should be closely related to the similarity of the risks. Finally there is abundant evidence in social psychology (Isen, Shalker, Clark, and Karp 1978) for more pervasive influences of affect. We might expect that increases in estimated frequency might occur for all risks, a possiblity we term global generalization.

In the first two studies we examined the generalization of negative affect across the responses. Despite our attempts to provide a sensitive test of local or gradient generalization, both studies demonstrate sizable global generalization. Readers of the newspaper stories estimated that all causes of death were about 40% more common than the control. Since the changes were unrelated to the topic of the story, these data suggested that the effect was due to mood induced, and that the bad moods were more than unpleasant states. In addition, they had pervasive influences on an important class of risk-related judgments.

In the third experiment, we broadened the estimates we requested to include items not related to either death or risk. For example, subjects were asked to report the frequency of bankruptcy and divorce. Even with these widely divergent estimates, we have found strong global generalization of affect, with no evidence for either local or gradient generalization. We also included a condition which read an additional newspaper story free of risk related content, but which described a series of negative events which occurred to the main character. Since the story made no reference to risk or death, its principle effect was the negative mood it induced in the reader. This depressing story resulted in a pattern of results almost identical to those induced by the risk-related newspaper stories.

These data, viewed as a whole, demonstrate that affect can have a large and pervasive influence on one important class of judgments, estimates of the frequency of risk-related events in the population. So far we have found no indication of a connection between the information contained in a story and its impact on the estimated frequency of death. The overriding factor in these increases does not appear to be the story told, but rather the mood or affect state it conveys to the reader. These effects are not limited to areas of death, but have been shown for estimates of non-fatal hazards and lifestyle threatening risks such as divorce and bankruptcy.

Any model of affect must account for two important aspects of this phenomenon: (1) Induction of a negative mood alone is sufficient to change estimates, and (2) the size of the change is unrelated to the semantic similarity, either among the estimates themselves, or between the cause of the mood and the estimates.

References

Isen, A. M., Shalker, T.E., Clark, M. & Karp, L. "Affect, accessibility of material in memory, and behavior: A cognitive loop?", Journal of Personality and Social Psychology, 1978, 37, 1-14.