is apparent, on this basis, that xenon fluorides should be stable, and we may even predict that the enthalpy of dissociation of XeF_2 to Xe + F_2 lies between the ΔH values given above for the reactions of BrF_3 and ClF_3. Similarly, we may expect the enthalpy of the reaction $XeF_4 = Xe + 2F_2$ to be somewhat less than the 4.6 ev found (*3*) for $BrF_5 = BrF + 2F_2$.

Since the ionization potential of krypton is about 1 ev higher than that of chlorine, and since ΔH for the reaction $ClF_3 = ClF + F_2$ is only 1.1 ev, the ΔH of dissociation of any krypton fluoride is expected to be about zero, at best. Also, the entropy of dissociation of any complex fluoride will be positive, hence krypton fluorides are predicted to be thermodynamically unstable under low-pressure conditions, but might be prepared by indirect methods.

The corresponding energy expression for the five atom units

$$\begin{array}{ccc} & F & \\ F & X & F \\ & F & \end{array}$$

is essentially similar to Eq. 1 for F—X—F, but it involves also the second ionization potential of X and a larger ionic bonding term as well as a coefficient of 2 for the other terms in Eq. 1. Detailed calculations indicate that, for the elements under discussion, the boundary of stability is essentially the same for the XF_4 unit and for the XF_2 unit. The F — F nonbonded exchange repulsions would be relatively more numerous in the XF_4 unit and larger, because of the shorter distances, in chlorine fluorides than in the other cases; this may account for the absence of a stable ClF_5.

The bonding in compounds such as ClO_2 or ClO_4^- is also partially ionic and partially covalent. Consequently the ionization potential of the central atom is expected to be a good indicator of stability of oxides of this type. Since the ionization potential of xenon is less than that of chlorine, attempts to produce xenon oxides seem to be indicated and are being made in this laboratory. Xenon oxyfluorides have been reported (*1*). It should be noted that most of the oxy-chlorine compounds are thermodynamically unstable with respect to the evolution of oxygen, whereas ClF_3 is stable with respect to the evolution of fluorine. Thus, indirect methods of preparation may be required for xenon oxides, and such compounds, even if formed, may be unstable and possibly explosive.

KENNETH S. PITZER
Rice University, Houston, Texas

References and Notes

1. C. L. Chernick *et al.*, *Science* **138**, 136 (1962); H. H. Claassen, H. Selig, J. G. Malm, *J. Am. Chem. Soc.* **84**, 3593 (1962).
2. L. C. Allen, *Science* **138**, 892 (1962); ——— and W. D. Horrocks, Jr., *J. Am. Chem. Soc.* **84**, 4344 (1962).
3. Wiebenga, Havinga, Boswijk, *Advan. Inorg. Chem. Radiochem.* **3**, 133 (1962).
4. In the case of iodine this process goes even further, with fluorine, to yield IF_7; this last compound is less analogous to the xenon fluorides and is not considered further.
5. R. Livingston, *J. Phys. Chem.* **57**, 496 (1953); D. F. Smith, *ibid.* **21**, 609 (1953).
6. C. D. Cornwell and R. S. Yamasaki, *ibid.* **27**, 1060 (1957).
7. R. S. Yamasaki and C. D. Cornwell, *ibid.* **30**, 1265 (1959).
8. The ionic term is determined by the bond distances, which are nearly constant: Cl—F is 1.70 Å; Br—F is always near 1.81 Å; I—F in IF_5 is reported to be approximately 1.75 Å, but I—F is probably at least as long in IF_5 as in BrF_5. The covalent bond energies in the series ClF, BrF, IF are likewise nearly constant: 60.3, 59.4, and 66.2 kcal, respectively.

26 December 1962

Word Association: Common and Original Response

Abstract. *The scoring of popular responses in the word-association test provides only a fair estimate of the number of original responses. The magnitude of the relationship between popular and original responses will be overestimated unless one employs a correction for the constraint imposed by one score on the other. Two methods of correcting for the lack of independence between the numbers of popular and of original responses are described.*

In 1910, Kent and Rosanoff selected 100 common English words (none of them especially likely to provoke socially dubious responses) as stimuli; presented the words verbally to 1000 "normal men and women"; and recorded their verbal responses (*1*). Scoring the association-test results only for "individualistic" responses (that is, responses made only once), Kent and Rosanoff found that the percentages of individualistic responses for different groups were as follows: for "normal persons with only common school education," 5.2; for college-educated subjects, 9.3; and for some schizophrenic patients, 25 to 50. Individualistic responses have been of interest to psychologists studying many different aspects of thinking and problem solving. MacKinnon, studying creativity among architects (*2*), "found the unusualness of mental associations one of the best predictors of creativity."

Table 1. Correlation coefficients for the three indices of the relationship between popular and original responses.

Correction method	Combined ($N = 56$)	First test ($N = 28$)	Second test ($N = 28$)
Uncorrected	−.88		
Odd-even:			
Populars$_{odd}$ versus originals$_{even}$	−.74	−.76	−.71
Populars$_{even}$ versus originals$_{odd}$	−.74	−.75	−.80
Proportion	−.78	−.78	−.79

While investigators in a number of areas are interested in idiosyncratic, novel, improvisatory, or original behavior, it is obviously very time-consuming and laborious to score a test for rare responses. In the case of the stimulus word *Needle* in the word association test, one would have to run through from 45 to 60 responses in order to determine that the response being scored is truly of the one-in-a-thousand variety. Woodworth and Schlosberg assure us that the much easier task of scoring only the most common response to each stimulus word provides an inverse measure of rarity; however, they do not cite evidence in support of this assertion (*3*).

Woodworth may have reached that conclusion after seeing correlation coefficients computed from a comparison of gross numbers of popular and of original responses (*4*). Such a procedure would yield a spuriously large negative correlation, for the number of popular responses limits the number of original responses that are possible; one could not possibly give a large number of popular responses *and* a large number of original responses. One can correct for this lack of independence of the two measures by determining, for each subject, the proportion of responses that are popular responses and correlating that value with the proportion of the remaining responses that are original responses [that is, compute the correlation r between $P/100$ and $O/(100 - P)$]. Another method of correction resembles the split-half method of computing test reliability. It is conceptually different from the proportion-correction method, but the two yield similar estimates of the true relationship. Correlation coefficients obtained

Table 2. Reliability coefficients for original and popular responses computed by test-retest and odd-even methods.

Response	Combined ($N = 56$)	First test ($N = 28$)	Second test ($N = 28$)
Test-retest reliability estimates			
Popular	.80		
Original	.86		
Odd-even reliability coefficients			
Popular	.82	.85*	.81
Original	.91	.91	.87

* N = 46 for this sample.

by these three methods of computation for describing the relationship between numbers of popular and of original responses are given in Table 1. The difference between the coefficients obtained by the correction methods is not significant; both coefficients are significantly different from the uncorrected r ($p<.05$ and $p<.01$).

The raw data were obtained by administering the Kent-Rosanoff word association test to 28 volunteer college students on two days 15 days apart. The respondents did not realize that they would take the test a second time until they were presented with the task 15 days after first taking it. The instructions for the two presentations were identical—namely, to give the first single word that comes to mind after reading the stimulus word.

The results from two methods of estimating the reliability of popular-response and original-response scores are presented in Table 2. By either method of computation, the original-response scores are the more reliable.

In summary, the scoring of popular responses in the word-association test provides a fair estimate of the number of original responses, accounting for from 55 to 62 percent of the variance rather than for the 77 percent estimated from an uncorrected correlation of popular with original responses. The number of original responses is significantly more stable than the number of popular responses; this is not surprising in view of the fact that more than one response can be scored as original for any given stimulus word but only one response can be scored as popular (*5*).

FREDERICK R. FOSMIRE
H. EDWARD TRYK
Department of Psychology, University of Oregon, Eugene

References and Notes

1. G. H. Kent and A. J. Rosanoff, *Am. J. Insanity* **67**, 37 (1910).
2. D. W. MacKinnon, *Am. Psychologist* **17**, 491 (1962).
3. R. S. Woodworth and H. Schlosberg, *Experimental Psychology* (Holt, New York, rev. ed., 1954). Similar statements are found in R. S. Woodworth, *Experimental Psychology* (Holt, New York, 1938).
4. *Popular response*, in our terminology, signifies the most common associate of any stimulus word; *original response* signifies an associate not included in the 1954 University of Minnesota norms for 1008 college students [see W. A. Russell and J. J. Jenkins, "The Complete Minnesota Norms for Responses to 100 Words from the Kent-Rosanoff Word Association Test" (Univ. of Minnesota Press, Minneapolis, 1954)].
5. This research was supported in part by the National Institute of Mental Health [grants M-3841(A) and M-4978)]. Ralph Hollingworth assisted with the computations.

8 October 1962

Personality Test Interpretation by Digital Computer

Abstract. *In this study a set of decision rules was devised for interpreting profile patterns of the Minnesota Multiphasic Personality Inventory (MMPI) of maladjusted and adjusted college students. The procedure used was that of computer programming of the "maladjusted" versus "adjusted" decisions of an expert test interpreter. The interpreter's decision-making processes were tape-recorded while he was thinking aloud during the sorting of the profiles of 126 college students. The programmed decision rules, which were based on the interpreter's protocol and which were improved upon by a process of trial-and-error statistical checking, yielded a greater hit percentage than the decisions of the original interpreter. In its final form, the set of objective configural inventory rules identified correctly large numbers of maladjusted college students in two cross-validation samples.*

Within the past few years research on human thinking has been facilitated by the introduction of the electronic digital computer as a research tool in the behavioral sciences and by the demonstration that this tool is much more than just a machine which performs rapid arithmetical operations. Among those who have contributed most to this development have been Allen Newell and Herbert A. Simon of Carnegie Institute of Technology and J. C. Shaw of the Rand Corporation (*1*). These workers have provided considerable and impressive evidence that the digital computer, when appropriately programmed, can carry out complex patterns of processes.

We now report how a computer has recently been applied as a tool to aid in interpretations of personality tests. The computer was used to approximate the rules expressed in the tape-recorded verbalizations of an expert test interpreter.

One of the personality tests which has frequently been used in psychiatric settings to aid in diagnostic and prognostic decision-making is the Minnesota Multiphasic Personality Inventory. The inventory is conventionally scored on a profile sheet which contains four validity and ten clinical scales. In this study an experienced user of the inventory was instructed to discriminate between the test profiles for maladjusted college students ($N = 45$) and those for well-adjusted students ($N = 81$).

The profiles for 126 college students (72 males and 54 females) were used as a criterion sample upon which a set of decision rules was developed. Such profiles were obtained from students who belonged to subgroups as follows.

1) Adjusted and maladjusted counseling group. This group was comprised of 65 students who had voluntarily requested help from the Carnegie Institute of Technology Counseling Center. They were judged by two counselors, after the completion of several interviews, to have problems of either a vocational-academic or a personal-emotional nature. The students whose problems were vocational-academic were labeled "adjusted" ($N = 37$); those whose problems were personal-emotional were called "maladjusted" ($N = 28$).

2) Adjusted and maladjusted no-counseling group. There were 31 students in this group, all members of fraternities and sororities. They were classed as either "adjusted" or "maladjusted" on the basis of the way in which their fraternity brothers (or sorority sisters) perceived them. Each member of each fraternity and sorority on campus, under supervised conditions, nominated from a roster of names of his fraternity brothers or sorority sisters four individuals, two of whom he considered the least well adjusted and two the best adjusted. A student was retained in this group if 60 percent or more of his peers nominated him, or her, for one of the two categories. Finally, 31 of these students (the 17 least well adjusted and the 14 best adjusted) were given the personality inventory test. This type of sample was chosen in order to counterbalance the