Journal of Psycholinguistic Research, Vol. 1, No. 3, 1972

Factors of Style and Personality

Ernst Moerk[1]

Received July 8, 1971

In an attempt to find relationships between psychological and linguistic variables, style samples of short stories, 300 words each, were analyzed according to formal criteria and the results were correlated with scores on personality tests. The number of significant correlations supported the hypothesis that style is related to personality. A factor analysis, using the principal component solution and Varimax rotation (Harman, 1967), of the correlation matrix resulted in six identifiable factors of style, three factors of psychological tests, and a large number of small factors, each represented only by two to five experimental variables with significant loadings. A significant loading for interpretation purposes was defined in agreement with Guilford (1956) as 0.30 or greater, positive or negative. A close examination of style factors led to the tentative differentiation of a basic language factor as resulting from grammatical constraint, and several factors of subjective style of individuals. Some of these later factors had enough loadings on personality variables to permit cautious psychological interpretation.

INTRODUCTION

It is unquestioned by scientists of diverse provenience that there is a close connection between personality and language behavior of a person. A good deal of research has been done to define more closely these relationships, and this material is summarized and evaluated by Mahl and Schulze (1964). For the present paper, the main attention was directed toward the relationship between the style of writing of persons as measured through mainly formal characteristics as opposed to a study of content and a broad range of personality traits as measured by psychological tests.

To approach this problem, linguistic units have to be defined which are

[1]Fresno State College, Fresno, California.

psychologically meaningful and which can be objectively measured. Furthermore, psychological tests must be selected which measure traits relevant for language behavior. Both these tasks could only be fulfilled satisfactorily on the basis of a sound psycholinguistic theory concerning the relationship of language to personality. Such a theory does not yet exist in spite of great interest in psycholinguistics during the recent decades. Osgood and Sebeok (1965) point out that we have to consider multilevel units of style which may range in complexity from the single letter to phenomena which encompass more than one sentence or even entire paragraphs. Style criteria of different levels of complexity were therefore chosen for analysis. Although relationships of a few specific personality traits, such as anxiety, to stylistic aspects are demonstrated, the selection of a broader variety of tests for the present research had still to be purely inductive. It was not attempted to cover all dimensions of personality, but to get some first cues through the use of a considerable diversity of tests.

METHOD

Thirty freshmen college students of both sexes wrote short stories according to the Incomplete Story Method (Ungricht, 1955). The beginning sentence was given: "He (She—according to the sex of the subject) stood at the window, clasped his (her) hands behind his (her) back and stared out into the night." *S* was required to imagine a person of his own sex standing at the window but otherwise was completely free to continue the story however he wished. Through this method, two goals are approximated: to leave nearly complete freedom to choose the topic, but also to keep the content of the stories from becoming too diversified. Previous research (Antosch, 1949) has shown that different content categories, such as scientific writing *versus* fairy tales, influence the form as measured in style variables. As experience shows, this one sentence induces practically all writers to adhere to an area of content concerning personal problems, social interaction, feelings, and memories, so that content *per se* should produce only minimal differences in style or none.

In the course of a few days after the collection of this style sample, two personality tests, the MMPI and the Rorschach; an intelligence test, the Wechsler Adult Intelligence Scale; and an interest test, the Kuder Preference Record, Professional, were administered to each *S*. This battery provided the desired diversity in traits tested, as well as in testing methods used.

Only those style variables were chosen for analysis where inter-rater agreement was practically a hundred percent, to obtain the highest possible

objectivity. They ranged from the smallest unit, the letter, to variables concerning sentence structure and those exceeding the limits of the single sentence, like rhythm of nouns or perseverations. Word length was counted in letters, from words consisting of only one letter to words of the length of 10 or more letters. Words were distinguished by grammatical category and according to their function in the sentence. The following groups were selected: preposition, noun as subject, noun after preposition, noun as object, helping verb, main verb without "to be," verb "to be," personal pronoun, possessive pronoun, coordinating conjunction, subordinating conjunction, adjective used attributive, adjective used predicative, adverbs which can only be used as adverbs, like very, yet, always, perhaps; and adverbs derived from adjectives, like beautifully, hardly, etc. It was attempted to include sentence structure specifically through the following categories: the use of the passive form, intransitive verb, transitive verb, and infinitive. Sentence length was expressed in four measures: maximum number of words per sentence, minimum number of words per sentence, mean number of words per sentence, and variance of number of words per sentence. The four categories of definite article, indefinite article, frequency of persons mentioned, and frequency of the use of the neutrum differ from the previous categories in that they are at least partially variables of content and not only of form. It was hypothesized that their correlations with personality variables could be different, as they combine syntactic and semantic functions. Criteria going beyond the boundary of the single sentence are the total number of words used in the essay; the rhythm of nouns, defined by counting how often nouns followed each other with a distance of 0-1 up to 8-9 words; and the rhythm of perseverations, found by counting how often the same word is repeated with a distance of 0-1 up to 8-9 words. These single measures were summarized into two large groups: repetitions with distances of 0-4 words, and with distances of 5-9 words. The basis for all these counts was the first 300 words of the story; only the total number of words was counted from the entire essay.

Product-moment correlations were calculated from the data. Three kinds of correlations between scores resulted: (1) Correlations between style and personality traits. According to the null hypothesis, writing style would only be dependent on the laws of language and would not be related to specific personality traits. The expected correlations would, therefore, be close to zero. (2) Correlations between style variables which should be, on a purely logical basis, slightly negative, as in a finite sample of words each appearance of one word or form makes the appearance of any other word or form less probable. (3) Correlations between personality scores which are, as experience shows, often positive. These last correlations are immaterial to the present investigation and are therefore not discussed further.

The correlation matrix was subjected to a factor analysis using the principal component solution and Varimax rotation (Harman, 1967). Factors were extracted until 99% of the total variance was covered, but only the first eight factors, covering approximately 70% of the variance, could be tentatively interpreted, as all the other factors were represented by only two to five experimental variables with significant loadings. It is fully recognized that with an N of 30 subjects, many correlations which may express an actual relationship will be insignificant. The low number of subjects has especially adverse effects as the correlations between style and personality are expected to be low because of the multidetermination of single-style phenomena. It represents further a serious obstacle for the interpretation of the resulting factors. But the design had to follow the basic principle of cost-reward balance. In a completely uncharted field like the one investigated, the results, even with a large N, would be very tentative as they depend upon the style variables and tests selected. The costs with a large N would, however, be so high as to be prohibitive. The choice for the investigator lies, therefore, at present between complete neglect of the topic or a tentative and statistically hazardous search for ideas and suggestions. The author chose the last alternative.

RESULTS

The present paper will contain mainly the discussion of the correlations between style and personality traits. Fifty-one style variables were correlated with 64 scores of the psychological tests. The resulting correlations were relatively low and, in a high percentage, insignificant. On the other hand, the patterns of significant correlations are revealing.

Sixteen of the 64 variables of psychological tests showed more significant correlations with style characteristics than could be anticipated if only chance connections existed. As Table I shows, seven of the Rorschach variables are related to personal style, with the quotient (H+A)/(Hd+Ad), the sum of whole-human plus whole-animal answers divided through the sum of human- plus animal-part answers, reaching a peak of 15 significant correlations. The professional interests of a person seem also to be of importance for his style of writing as evident through the significant correlations of the Kuder Preference Test. The scales for art and music, with 12 and 10 significant correlations respectively, especially stand out. In contrast to both these tests, only one scale of the MMPI, hysteria, with nine correlations, is more closely related to style. Even intelligence, as measured through the WAIS, shows only a moderate number of significant correlations, with total IQ reaching the

Table I. Correlations of Psychological Tests with Style Variables[d]

Test	Subtest score	N of significant correlations	Magnitude of correlations[a]		
			$P = 0.05$	$P = 0.01$	$P = 0.001$
MMPI	Hysteria	9[b]	–43,–38,37,–44, –41,–42,–38	–54,–49	
Kuder	Art	12	–43,–38,–41,–46, –44,–39,–45	–47,–51, –50,–49	–61
	Literature	6	–41,–43,39,39, 41	–48	
	Music	10	–36,–41,–39,40, –36,40,43,37	48,49	
WAIS	Information	7	39,39,38,36	48,47,48	
	Vocabulary	4	38,38,41,–38		
	Digit symbol	5	45,37,38,39	46	
	Verbal IQ	6	40,41,37,42,41	50	
	Total IQ	8[c]	37,39,44,39,38, 37,43,38		
Rorschach	W	5	39,37,–40	46,49	
	M.FM.m	4	–49,–40	–53	–60
	H+Hd	4	–37,–39,–38	–52	
	F/R	6	–40,37,46,39	–55	–62
	A+Ad/R	5	39,42,42	53,56	
	(H+A)/(Hd+Ad)	15	39,45,44,38,44, 43,38,40	56,50, 46,50	64,57, 70
	W/M	5	–37,–36,39,–41	48	

[a]Decimal points have been omitted.
[b]Of 51 correlations of each test score only around 2 to 3 would be expected to be significant at the 5% level through chance alone.
[c]As the value of IQ total consists to a great part of the value of the verbal IQ, this variable cannot be counted independent of the verbal IQ!
[d]Standard error of $r = 0.185$.

highest number of eight correlations. Most of the selected style variables are not closely connected with personality.

But as a cue for future research, Table II displays evidence that more than half of the variables which have more significant correlations than expected by chance belong to a category best described as rhythm: word length, sentence length, several patterns of perseveration, and one kind of

Table II. Correlations of Style Variables with Tests[a]

Style variable	*N* of significant correlations	Magnitude of correlations[b]		
		$P = 0.05$	$P = 0.01$	$P = 0.001$
Preposition	8[a]	-43,-43,-40,43 -36,39,37	48	
Noun as object	8	40,39,37,37,45	51,50,-47	
Helping verb	6	40,39,-36,39, 39,39		
Verb "to be"	8	44,39,37,-41,40, 43,44,40		
Personal pronoun	8	37,-39,38,41,43, 40,41,39		
Infinitive	8	45,38,37,40,37, -40,-37	51	
Frequency of the use of neutrum	10	43,40,37,41,45, 45,44,42	48,-49	
Words with four letters	7	42,41,-36,38	48,49,47	
Words with nine letters	6	-42,-44,-40, -41,36	-47	
Maximum number of words per sentence	6	-46,-42,38,40	-54	69
Weighted number of perseverations	6	-41,38,-39,37	-53,53	
0-4 words between perseveration	5	-36,-37,-38,39	-48	
5-9 words between perseveration	5	-39,-36,46	56	-60
2-3 words between perseveration	5[c]	-43,-40,37,39,42		
4-5 words between perseveration	5	-37,-37,42	-49,-51	
6-7 words between nouns	7	36,-36,48,-38, -39	49,46	

[a]Of 64 correlations of each style variable only around 3 to 4 would be expected to be significant at the 5% level through chance alone. This number would be even lower at the 1% or 0.01% level of significance.

[b]Standard error of $r = 185$. Decimal points have been omitted.

[c]As the number of 2-3 and 4-5 words between perseverations is included in the number of 0-4 and 5-9 words between correlations, only 14 variables can be counted as significantly correlated to personality characteristics.

noun patterns are all phenomena of the flow or movement of writing. A similar relationship of psychological and rhythmic variables was established by Mittenecker (1951) who found correlations between the phenomena of perseveration in speech and the psychological disturbance of schizophrenia. Word length as style-specific phenomenon was described by Fucks (1959). Grammatical structure proved to be little connected with psychological variables as evident in Table II. Only 6 of the 19 measured grammatical variables show a moderate number of significant correlations. Neither seem semantic components in some of our criteria to influence their relatedness to psychological variables. Only one of the partially semantic variables, frequency of the use of neutrum, has a closer affinity to personality, with 10 significant correlations.

From the purely linguistic point of view, it would be interesting to discuss intensively the pattern of correlations between style variables themselves. These correlations and the derived factors may be seen as expression of two basic principles: (1) the influence of the laws of language structure, and (2) psychological characteristics, e.g., when a person who uses long words also builds long sentences. As the present investigation was mainly psycholinguistically oriented, the purely linguistic aspect will be pointed out only briefly in the discussion of the factor analysis.

The factor analysis produced only a few factors with high loadings in style and psychological variables. Further factors of special psychological tests resulted, like a "Rorschach-factor," a "WAIS-factor," and a "MMPI-factor." The scales of the Kuder Preference Record, however, are split up in many factors, so that practically every subscale forms its own factor, together with a few loadings from style and other personality tests. And it is evident from Table III that many of the subtests or subresults of the aforementioned personality tests, Rorschach, WAIS and MMPI, contribute to the structures of other factors. About 15 factors of style also show loadings in variables of personality tests. Only six of them have enough substantial loadings that a description can be attempted. Table III summarizes these factors, the variables contributing to them, and the loadings of these variables.

Factor one has its highest loadings on the basic variables of language, like number of words in the essay, or number of main verbs, with many loadings in the range of 0.8 and 0.9, and it seems, therefore, to describe the basic style as determined by general laws of language. Only a few variables of the personality tests contribute to this factor, and even they have only low loadings, so that an interpretation in terms of personality characteristics is impossible at present.

Factor two shows no loadings at all from variables of personality tests, and also seems to be very narrow even in regard to style phenomena. It could

Table III. Factors of Style

Factor 1 Basic style		Factor 2 Noun style		Factor 3 Persevering style		Factor 4 Precise style		Factor 5 Enriched style		Factor 6 Rhythm of style	
Total *N* words	0.976	0-4 words between noun	0.957	Weighted number of perseverations	0.948	Variance words/sentence	−0.941	Adverb from adjective	0.800	6-7 words between nouns	−0.883
Main verb	0.960	Weighted number of noun repetitions	0.943	0-4 words tween perseverations	0.884	Maximum words/sentence	−0.873	Passive	0.710	5-9 words between nouns	−0.550
Intransitive verb	0.885	2-3 words between nouns	0.689	2-3 words between perseverations	0.848	10+ letters/word	−0.759	8-9 words between nouns	0.595	Passive	0.396
Adverb	0.858	0-1 words between nouns	0.675	5-9 words between perseverations	0.783	10 letters/word	−0.743	5-9 words between nouns	0.562	4-5 words between nouns	0.346
Frequency of persons mentioned	0.854	4-5 words between nouns	0.636	4-5 words between perseverations	0.669	Transitive verb	−0.333	1 letter/word	−0.385	1 letter/word	0.335
Personal pronoun	0.839	Noun after preposition	0.499	6-7 words between perseverations	0.610	4-5 words between nouns	−0.324	Possessive pronoun	0.359	9 letter/word	0.325
"To be"	0.823	Preposition	0.411	Possessive pronoun	0.425	Adjective attributive	−0.302			5 letter/word	−0.300
Frequency of neutrum	0.800	Noun-object	0.368	9 letter/words	0.373	5-9 words between noun	0.302	Rorschach:			
Helping verb	0.797	Frequency of neutrum	0.328	8-9 words between perseverations	0.316	Coordinative conjunction	−0.301	R 8-10 R	−0.877	WAIS:	
Noun as subject	0.797	Definite article	0.318	Frequency of persons mentioned	0.310	Rorschach:				Arithmetic	−0.344
Preposition	0.778	Possessive pronoun	0.313			M/Sum C	−0.720			Vocabulary	0.312
Coordinative conjunction	0.702					F/R	0.639				
Transitive verb	0.663					MMPI:					
Definite article	0.660					Hysteria	0.329				
Noun-object	0.618										
Noun after preposition	0.593										
Infinitive	0.584										

Indefinite article	0.581
Adjective attributive	0.570
Adjective predicative	0.380
Possessive pronoun	0.378
Mean words/sentence	−0.365
Adverb form adjective	0.362
Minimum word/sentence	−0.328
Subordinate conjunction	0.325
4-5 words between perseveration	−0.314
Rorschach:	
(H+A)/(Hd+Ad)	0.425
Original	0.409
WAIS:	
Digit symbol	0.306
MMPI:	
Hysteria	−0.328
Psychopathic deviance	−0.301
Kuder preference:	
Mechanic	0.307
Rorschach:	
A+Ad/R	0.447
M,Fm,m	−0.341
F/R	0.341
WAIS:	
Digit symbol	0.411
MMPI:	
Paranonia	−0.336
Kuder preference:	
Literature	−0.316
Kuder preference:	
Art	0.339

be described as a factor of a "noun style." The variables with loadings in this factor are either nouns in different forms of grammatical use, or words which are closely connected with nouns through the laws of grammar, like article, and preposition. But the highest loadings are found in variables which were selected to measure rhythm of nouns. The factor, therefore, describes the noun dimension, with many nouns which consequently follow upon each other in short intervals.

Factor three has several high loadings in the variables of perseveration, which were chosen according to the hypothesis that they are related to specific personality structures. The weighted number of perseverations, which encompasses perseverations over very small and larger intervals, with highest weights attributed to the perseverations over small intervals, has the highest loading of 0.948. Furthermore, nearly all perseverations as measured according to the distances between repeated words have appreciable loadings in the range of 0.32 to 0.88. The variable possessive pronouns showing up in this factor is probably also only one form of perseveration, as the identical pronoun can be easily repeated after small intervals. In addition to these style variables, factor three includes loadings in variables of several personality tests, as seen in Table III. A cautious interpretation of these variables gives the picture of an unimaginative person (A+Ad/R = 0.447) restricted in his emotional experiences (M,FM,m = -0.341, literature = -0.316) who can concentrate on his work (digit symbol = 0.411, F/R = 3.41) but is not creative or imaginative. This picture as evolving from the test variables corresponds closely to Mittenecker's description of schizophrenic patients where he found this style and is also in congruence with intuitive expectations concerning a person who is quite repetitious. This factor comes consequently closest, of those discussed until now, to being really a factor of style and personality, where both contributing sides can be clearly discerned. It is expected that future research will add several more variables of both categories with substantial loadings.

Factor four is quite clearly definable as denoting two opposites of style, namely short precise *versus* lengthy prose. This continuum is expressed in sentence length as well as word length: the loadings of variation of words per sentence -0.941; maximum words per sentence -0.873; words with more than 10 letters -0.759; and words with 10 letters -0.743, are high and significant. The person writing precisely has a high number of form responses in the Rorschach (F/R = 0.639), a low value of the ratio M/Sum C (loading -0.720). He is more interested in art (loading 0.339) and the variable hysteria of the MMPI carries a moderate loading of 0.329. While by far not psychologically unambiguous, factor four seems to describe a clear entity of style and personality.

A fifth factor still shows several variables with high loadings. Yet a verbal characterization is already more difficult. The loadings of adverb derived from adjectives (0.800); passive voice (0.710); 8-9 words between nouns (0.595), and one-letter words (-0.385) designate the factor as a dimension of enrichment. A single high loading on a personality variable does not contribute to the elucidation of the factor.

The last factor has enough significant loadings to prove its statistical reality, but linguistic or psychological interpretation is nearly impossible. According to the two highest loadings, it could come close to expressing the phenomenon of rhythmicity, but this description must remain very tentative.

DISCUSSION

By reason of the small number of subjects, the present results can only be viewed as very tentative and preliminary. While the basic hypothesis of a connection between personality characteristics and style phenomena was supported through a number of significant correlations, the magnitude and number of these correlations was comparatively low. This result may be attributable to the specific variables selected, to the low number of *S*s, or to a genuine relative independence of specific style and personality traits. The large number of 15 style factors, most of which have only a few significant loadings, creates the impression that the individual style seems to be a function of many comparatively unrelated mechanisms. A glance at the style variables, which show a more than accidental number of significant correlations, makes it clear that these variables cover a broad field of style characteristics and that consequently personality traits are involved in many respects in producing the individual style. A more thorough sampling of personality measures and style criteria is needed for the attainment of a more complete picture.

The lack of a theoretical basis for the selection of variables has been stressed. Empirical evidence from this study seems to point more toward the creative, emotional, projective side of personality traits, where relationships between style and personality could be found. The range of style variables correlated with personality traits is so wide and diversified that no clear pattern can be discerned at present.

REFERENCES

Antosch, F. (1949). Literatur und Sprachuntersuchungen mit dem vereinfachten Aktionsquotienten. Unpublished doctoral dissertation, University of Vienna.

Balken, E. R., and Masserman, J. H. (1940). The language of phantasy. *J. Psychol.* 10: 75-86.

Busemann, A. (1926). Ueber typische und phasische Unterschiede der kategorialen Sprachform. *Z. Paedagogische Pyschol.* 27: 415-427.

Cattell, R. B. (1965). *The Scientific Analysis of Personality.* Penguin Books, Baltimore, Md.

Fucks, W. (1952). On mathematical analysis of style. *Biometrika* 39: 122-129.

Guilford, J. P., and Zimmerman, W. S. (1956). Fourteen dimensions of temperament. *Psychol. Monographs* 70 (10, Whole No. 417).

Harman, H. H. (1967). *Modern Factor Analysis.* (2nd ed.), Unviersity of Chicago Press, Chicago, Ill.

Mahl, G. F., and Schulze, G. (1964). Psychological research in the extralinguistic area. In Th. A. Sebeok (ed.), *Approaches to Semiotics: Cultural Anthropology, Education, Linguistics, Psychiatry, Psychology,* Mouton, The Hague.

Mittenecker, E. (1951). Eine neue quantitative Methode in der Sprachanalyse und ihre Anwendung bei Schizophrenen. *Monat. Psychiat. Neurol.* 121: 6-21.

Osgood, Ch. E., and Sebeok, Th. A. (1965). *Psycholinguistics: A Survey of Theory and Research Problems.* Indiana Univ. Press, Bloomington, Ind.

Sanford, F. H. (1942*a*). Speech and personality. *Psychol. Bull.* 39: 811-845.

Sanford, F. H. (1942*b*). Speech and personality, a comparative case study. *Character and Personality* 10: 169-198.

Ungricht, J. (1955). Der "Sohn-Aufsatz" als psychodiagnostisches Hilfsmittel. *Schweizerische Z. Psychol.* 14: 27-45.