Classics in the History of Psychology

(Return to *Classics* index)

A Study of Mental Statistics

Joseph Jastrow (1891)

First published in *New Review*, *5*, 559-568.

Posted August 2000

To anyone impressed with the importance of objective results it is always gratifying to throw into numerical form the result of subjective or partially unconscious operations. The difficulties of introspective observations are many and obvious; such observations are warped, not only by the thought-habits of the individual observer, but they labour more particularly with the difficulty that if we allow our mental processes to go on in their natural trend, our memory of them soon fades away and becomes distorted, while in so far as we turn about to stare at them as they pass through the mind their original purity passes from them and leaves them artificial; they are like children romping about unconcernedly and expressing themselves freely in the privacy of the family circle, but bashful, silent, and conventional before strangers.

In the hopes of obtaining material capable of yielding in objective form the results of unconscious and natural mental processes, I asked the students of my class in Psychology to write out at their leisure one hundred words as rapidly as possible and to record the time. More definite instructions were purposely omitted, save that a caution against writing words in sentences was added. In this study I shall confine my results to those obtained from the papers of all the lady students in my class (twenty-five) and those of the twenty-five male students who happened to hand in their papers earliest. This material offer an embarrassing number of points of interest, amongst which I shall restrict myself to the following three: (*a*) the Community of Ideas and Thought-habits; (*b*) the Nature of the more usual Types of Association; and (*c*) the Time-relations of these Processes.

(*a*) In intellectual matters we are prone to insist upon our own [p. 560] individuality; we regard our thoughts as peculiarly and originally ours, and so easily tend to look upon any similarity of thought as extremely surprising. It is this tendency that magnifies all such evidences of mental community into startling coincidences. Moreover, the habit is likely to remain unchecked because so many of these similarities are founded in the region of the unconscious, and thus are overlooked by the testimony of conscious memory. None the less we are (or rather ought to be) guided by the same logic; and it is the aim of modern Psychology to show the same laws at work in the minds of all. Even in the most recondite walks of science we find discoveries made independently by different observers, and unconscious plagiarism is a well-recognised occurrence. "Public Opinion" and it Spirit of the Age" are formed, to a great extent, by unconscious similarities of thought. The probable extent of such similarity depends upon the usualness of the mental task in question. In the present case the mind is put upon its mettle to work as speedily as possible, and it naturally draws upon the most familiar land accessible shelves of its storehouse. Accordingly I find that of the entire 5,000 words only 2,024 are different. This means that fifty persons independently writing any one hundred words from the large number with which they are familiar, all in all, select from the same limited number of some 2,000 words -- certainly a striking illustration of the limitations of the everyday workings of the mind and the community of our

interests. Perhaps the result becomes still more striking: if we take into account the words occurring but once in the entire lists; of these there are 1,266, and subtracting this number from the 5,000 words as well as from the 2,024 different words, we find that the remaining 2,976 are formed by the repetitions of a handful of 758 words. Still further, if we take the one hundred words most frequently mentioned, we find that the Sum of their occurrences is 1,504, or three-tenths of the entire list.

Proceeding to the analysis of this "mental community," there are three points of special interest: (*a*) Is this tendency to repetition greater in any portion of the hundred words than in the rest? (*b*) Is it greater in any special classes of words? (*c*) Are [p. 561] there any characteristic differences between the two sexes in this respect? Under the first lead there is a variety of evidence showing that the most common words are written at the head of the list, resource being had to the less usual words as the former give out. As the evidence for this statement is rather complicated, it may be sufficient to mention that in the three groups of 500 words each, formed by the ten words at the head, the ten in the middle and the ten at the end of each list, the 25 words most frequently mentioned[1] occur respectively 154, 47, and 45 times; or 90 times oftener, 17 and 19 times less often than would result from a chance distribution.[2]·

In order to discuss the second point it becomes necessary to refer to the table here appended. In the first column we find 25 classes in the order of the frequency of their occurrence. It may suffice at present to notice that these classes are very unequal in size, purely empirical in origin, and serve to show the kinds of concepts most apt to occur, and to be represented by the same words in the minds of different persons. In the succeeding columns we find the *number of words*; of *different* words, and of *unique* or once-mentioned words for each of these classes distinguishing the words of the *men's* lists from those of the *women's* [p. 562] lists. A direct comparison of the ratio of repetition of these various classes is rendered unfair by their inequality of size,[3]

TABLE OF FREQUENCY OF MENTION OF VARIOUS CLASSES OF WORDS, &C.

Class	Number of Words in each Class		Number of different Words			Number of "unique" Words	
	Male	Female	Male	Female	Total	Male	Female
1. Animal Kingdom ...	254	178	81	52	97	36	14
2. Wearing Apparel and Fabrics	129	224	48	63	78	11	23
3. Proper Names ...	194	153	151	126	249	106	89
4. Verbs	197	134	134	77	178	85	37
5. Implements and Utensils	169	121	99	53	120	51	15
6. Interior Furnishings	89	190	39	52	66	13	12
7. Adjectives	177	102	120	65	155	80	33
8. Foods	53	179	38	67	83	13	25
9. Vegetable Kingdom	121	110	55	43	80	30	19
10. Abstract Terms ...	131	97	106	71	161	81	48
11. Buildings and Building Materials ...	105	117	52	44	70	23	12
12. Parts of Body ...	101	105	40	36	53	14	11
13. Miscellaneous ...	91	97	69	68	119	46	43
14. Geographical and Landscape Features	97	80	38	33	53	10	13
15. Mineral Kingdom ...	74	96	26	32	39	6	12
16. Meteorological and Astronomical ...	85	76	35	29	43	10	5
17. Stationery	60	86	17	20	24	2	4
18. Occupations and Callings	71	47	52	38	82	36	26
19. Conveyances ...	62	52	31	29	44	12	13
20. Educational	34	76	11	28	32	2	17
21. Other Parts of Speech	96	5	59	4	62	41	2
22. Arts	33	61	19	30	38	6	14
23. Amusements ...	30	53	22	34	50	12	21
24. Mercantile Terms ...	30	29	24	15	35	18	9
25. Kinship	17	32	10	9	13	2	3
Totals	2,500	2,500	1,376	1,123	2,024	746	520

[p. 563]

but the conclusion (again resting upon somewhat complicated evidence) seems clear that the relatively greatest repetition occurs in the classes of least extent; for example in the limited class of *stationery*, the 146 words mentioned are made up by the repetition of the same 24 words, and but six words are "unique."

Rather than overload this paper with numerical results, I prefer to print the table in sufficient fulness[*sic*] to enable the reader to deduce from it a number of interesting points and proceed to the last point, the comparison of the male with the female students, regarding this mental community. The inference is unmistakable that the women repeat one another's words much more than the men.

The female students use only 1,123 (44.9 per cent.) different words, while their male class-mates use 1,375 (55.0 per cent.). Again, the former have written but 520 (20.8 per cent.) *unique* words, while the latter have written 746 (29.8 per cent.). With regard to classes the women show different preferences from the men. The repetitions in names of articles of apparel, of interior furnishings, preponderate with the women, but they have a much larger number of such words altogether; similarly, the men have more repetitions in the names of *animals* and mention more such names.

Professor Minot reaches the same conclusion that "Women's repetitions greatly preponderate "; this is not equally true of all classes of diagrams, but that in general "there is much less variety among women than among men."

(*b*) It will be well to introduce the study of the kinds of association by a few remarks upon the nature and size of the 25 classes. These classes were derived merely by observing what prominent groups of words occurred in the lists. The purpose was to have as few classes as possible, provided they were natural, were distinctive and as unambiguous as the nature of the case would allow. In cases of doubt the group of associations in which the word occurred was generally sufficient to decide its classification. The relative sizes of these classes may thus be said to indicate in a rough way the contents of the most accessible and best stocked shelves of the average mind. Foremost among these are the names of *animals*, and particularly of the domestic animals, the five most common of which -- *horse, cow, dog, cat,* and *sheep* -- are mentioned [p. 564] in all 118 times. (I have included here the words *boy, girl*, *man, woman,* which four words occur in all 118 times, because these words were usually mentioned as, and along with, the names of animals.) Next comes *apparel* -- things with which everyone is constantly surrounded, which we have to manipulate daily and use as an aid to the distinction between man and man. · Omitting the three merely verbal classes, proper names, verbs, and adjectives, we find the next largest classes to be implements and utensils, interior furnishings; and articles of food. About one-third of all the words comes under these five categories. Referring again to the table for further illustrations I will select for special consideration the very interesting preferences of the sexes for different classes of words. The class to which the women contribute most largely is that of *Articles of Dress, one word in every eleven* belonging to this class; while the men have 129 such words, the women have 224. The inference from this that dress is the predominant category of the feminine (or the privy feminine) mind is valid with proper reservations; but we should remember that the dress of a woman is more conspicuous, more complex, and more various than that of a· man, and that she has more to do with the making of it. The prevalence of words denoting the common articles of furniture and the interior fittings of a house -- the peculiar field woman's household instincts -- is quite as marked, such objects being mentioned 190 times by women and but 89 times by men. In *Foods,* the disproportion is even greater, the women mentioning words of this class 179 times and the men 53. How far this difference can be accounted for by a more intimate acquaintance with the processes of preparing food I will leave an open question. In words referring to the Amusements, Arts, and Educational Matters the women show an excess over the men which likewise points to a characteristic difference in the interests of the two. In the young men we find marked predominances for the names of Animals, of Implements, and Utensils, the names of Professions and similar relations, and very markedly for Abstract Terms. In general we may conclude that the feminine traits revealed by this study are an attention to the immediate surroundings, to the finished product [p. 565] to the ornamental, the individual, and the concrete; while the masculine preference is for the more remote, the constructive, the useful, the general, and the abstract.

The study of the kinds of association revealed by these lists proves to be a rather difficult task. In reading them over it was usually easy to see the connection between word and word; but to describe and classify this connection without doing it violence is a difficult task. In the present unsatisfactory condition of the psychology of association, I can do little more than describe the more prominent and interesting types here represented. Perhaps the nest obvious and most frequent is the association by natural kinds, *i.e.*, by such related objects as would in most cases come under some one of the 35 classes of our table. One list, for example, begins with *cat, dog, hen, chicken, pug, bird*, *cow, bull, heifer, calf, kitten, girl, dress, ribbon, boy, hat, boots, shoes, stocking, cap, necktie, slippers, pants,* &c., composed, as is seen, largely of names of animals and articles of dress. Another begins: *Hand, face, lip, teeth, tongue, palate, nose, eyes, forehead, hair, nail, foot, toe, fingers, body, chest, legs, knees, calf, cow, horse, pig, donkey, mule, mare, boy, man,* &c., the transition from Parts of Body to Animals proceeding by the double meaning of the word *calf:* An unusual series is: *Stars, Mars, Mercury, Venus; Jupiter, Saturn Uranus, moon, earth, equator, cancer, capricorn*. A feminine series is: *Bread, butter, meat, potato, turnip, cabbage, sugar, tea, coffee, milk, cream, mustard, catsup, pepper, salt, vinegar*.

A striking type of association is by the *sound* of the word. *Cow* suggests *calf*, not alone for the same reason that it suggests *ox*, but in addition because the initial sound is the same in both words. That this is more a matter of sound than or sight is indicated by such instances as *nail*, suggesting *knife, hair*, *hare*; *soul, sole*; or again in the series *learn, burn, earn*. With some this habit is entirely the dominant one, and these are probably what I have: termed "ear-minded" persons.[4] One man presents an extreme case of this, for out of his hundred words only 15 do not rhyme [p. 566] with one or other of the neighbouring words. Here are samples from his list:--

Run	Tear	Wallow	Rope	Man	Money	Hurty
Fun	Hair	Swallow	Pope	Can	Funny	Curry
Pun	Care	Tallow	Soap	Tan	Sunny	Flurry
Gun	Chair	Mallow	Elope	Pan		
Ton	Stair	Shallow				
Son	Fair					
	Flare					

Table	Relate	Rich	Latin
Stable	Dilate	Pitch	Satin
Cable	Expatiate	Stitch	
	Dictate	Itch	

This rhyming is quite usual, and especially amongst the men. At times it is a mere similarity of sound, such as *picture* and *pitcher*, *cable*, and *cabal*, *butter*, and *letter*. An allied and prevalent tendency is that towards alliteration. I have counted how often each of the 25 most distinctive initial sounds is followed by word beginning .with the same sound, and find 390 such cases -- 240 amongst the men, 150 amongst the women. If this were a matter of chance there should only have been 200 of such cases altogether, 100 for each sex. To cite an extreme instance, one man, 75 of whose hundred words show this tendency, writes:-- *Hen, hand, harp, heard, head, heart, handy, hope, hop, hops, harpy*; and *ape, apron, apart, aspect, asp, Alps, air, arc, are*; and again, *finger, fin, fan, fond, fund, fancy*.

Other characteristic associations refer to such relations as: *Piano, Player* ; *musician, music; farm, farmer; student; professor; man, son; preacher, sermon*; and the like; to the passage from the general to the special, and *vice versâ*, such as *animal, horse; merchant, man; leaf, green; black, coal*; to the passage from whole to part, and *vice versâ*, such as *boat, oar,* and *handle, hammer*; synonyms, such as *hit, strike; tool, implement;* to parts of objects usually occurring in groups, such as *bread* and *butter, cuff* and *collar, knife* and *fork;* and to contrasts both of quantity and quality.

A further characteristic of these associations deserves notice. The mind does not always, and perhaps not often, proceed from [p. 567] one idea to another, and then to one suggested by the second, and so on, but one idea forms the centre, about which a small group of ideas accumulates, the mind in each case going back to the centre and out along a new radius. The simplest -case would be the alternation of associations such as occurs in the series just noted, where "*girl*" suggests "*boy*" across the two associations "*dress*" and "*ribbon*," while they in turn arch over the word "*boy*" to connect with "*hat*," and the idea of "*hat*" is probably still in the mind when "*cap*" is written four words later.

Finally, combining the two inquiries (*a* and *b*), I will investigate the similarity of the same word with the same word. For this purpose the 20 most frequently occurring words, whose aggregate occurrences number some 540 were selected, and in each case the preceding and the succeeding word were noted. In 30 cases one of these words began the entire list (proving clearly that the commonest words are apt to be first thought of), and in one case the word closed the list. The result was that of the 510 "preceding" words 111 (21.8 per cent.) were the same, and of the 539 "succeeding" words 145 (27.0 per cent.) were the same -- a very striking degree of community of association. It means that in the case of these 20 most common words the chances are about even that one or other of the neighbouring words will be the same as a designated associated word. I find, too that the chances of a word's suggesting the same word are somewhat greater than of its having been suggested by the same word. If we observe separately the results of the "feminine" and the "masculine" papers on this point, we find in the former 48 repetitions in 291 (16.5 per cent.) preceding words, and 55 (18.0 per cent.) repetitions in 306 succeeding words; and in the latter case 26 (11.9 per cent.) repetitions in 219 preceding words, and 40 repetitions (17.2 per cent.) in 233 succeeding words. As before, the women are much more like one another in their habits of thought than are the men.

(*c*) Regarding the time relations of these processes, allowance must be made for the somewhat rough methods of taking the time; as might be expected in so complicated and so varying a task the average variation is large. In the general average (in which these influences have a chance to correct one another) the writing and [p. 568] associating

of 100 words occupied : minutes and 8 seconds, or 3.08 seconds per word. · There is practically no difference in the times of the two sexes. The time includes the time of writing, of associating, and of doing both together; and inasmuch as writing is somewhat automatic in character the last element may be considerable. To determine this I found the time in a number of cases needed to write 100 disconnected words from dictation; this writing time proved to be 2.12 seconds per word, The difference between this and 3.08 seconds (0.98 seconds) is a *minimum* association time. Again, I had some: of the students call out as rapidly as possible 100 words in association. This takes 1.30 seconds per word, which may be regarded as a maximum association time; to read or speak these 100 words occupied 0.31 seconds per word, leaving 0.99 seconds as another *minimum* association time. From these data the average association time would be 1.17 seconds per word. We find, then (roughly speaking), that of the 308 seconds occupied in writing 100 words; 210 seconds are spent in the act of writing, 114 seconds in thinking of what to write, and 16 seconds in which both are done. ·

Footnotes

[1]

---

* It may be of interest to append a list of these words with their respective number of occurrences :—

1. Book 40	9. Cow		17. Pencil	
2. Horse 37	10. Paper } 26		18. Wood } 22	
3. Girl 35	11. Dress		19. Leaves 21	
4. Man... 34	12. Watch 24		20. Cat	
5. Boy 33	13. Hand		21. Pen } 20	
6. Table 30	14. Hair } 23		22. Hat	
7. Chair } 28	15. Dog		23. Shoe	
8. Tree }	16. House		24. Lake } 19	
			25. Knife	

Total number of occurrences 641

[2] Another method of testing the degree of repetition is by the percentage of words occurring but once in the entire list. Of these there are in all 1,266 or 25.3 per cent. In the two groups of 2,500 each, it is 29.8 per cent. male, and 20.8 per cent. female. For the three groups of 500 each the number of "unique" words (as I shall term them) is 216 for the "first" group, 315 for the "middle" group, and 343 for the "last" group, or in percentage 43.2, 60.0, and 68.6; the average of which is 58.3 per cent. The proportion of different words and the number of "unique" words are thus seen to follow parallel courses, both being indicative of the "community" of thought now under discussion.

[3] The ratio of repetition depends upon various factors, the influence of some of which may be briefly noted as follows:-- (*a*) *The number of persons writing the lists*, the indication being that the amount of repetition increases as this number increases. (*b*) *The number of words written*, affecting the repetition in the same way. (*c*) *The character of the task*, the inference being that the repetition increases as the task becomes simpler. Similarly the classes most largely represented will show a larger ratio of repetition; it seems probable, too, that, other things being equal, the repetition will be larger in classes of words that are homogeneous and familiar and in classes possessing a few very common representatives, not to mention other factors.

[4] See *Popular Science Monthly*, September, 1888.