

The Relation Between Psychological Change and Scientific Misconduct in Millikan's Oil-drop Experiments

Fumihito Ikeda*

Institute for the Advancement of Higher Education, Hokkaido University*

Abstract — Robert Andrews Millikan submitted two papers on measurements of the elementary charge of the electron to Physical Review in 1911 and 1913. In the 1911 paper, Millikan disputed the results of Felix Ehrenhaft. Then, in the 1913 one, Millikan manipulated the data to “enhance” the accuracy of his measurements. This manipulation could be described as scientific misconduct. In this paper, I examine the differences in wording between the two papers statistically and analyze the relation between the differences and the change in his mental state from the viewpoint of Transactional Analysis. In the results, there were more declarative expressions in the 1911 paper than in the 1913, and there were more expressions that emphasized the accuracy of measurements in the 1913 work than in the earlier one. From the viewpoint of Transactional Analysis, Millikan's mental situation changed from the “be strong” driver to the “be perfect” driver and this change led to his data manipulation. If the possibility of scientific misconduct could be detected in papers by this method, we could prevent such misconduct in the review process. In the future, I will examine the relations between wording and the drivers in Transactional Analysis more precisely and develop an automatic system detecting the drivers in scientific papers.

(Revised on 22 January, 2013)

Introduction

Although Robert Andrews Millikan (22nd March 1868 - 19th December 1953), who was a physician of U.S.A., was awarded the Nobel Prize in Physics at 1923 with studies on elementary charge and photoelectric effect, the former study, that the results had contributed to the 1911 paper (Millikan 1911) and the 1913 one (Millikan 1913) of Physical Review, had a story behind it: (1) Although the idea substituting oil drops for water drops to prevent from drying quickly had been proposed by Harvey Fletcher (11th September 1884 - 23rd July 1981) who had been under the direction of Millikan, Millikan convinced Fletcher to publish the result based

on this idea as his sole result on the condition that he permitted Fletcher to write his doctor thesis as sole author (Fletcher 1982). (2) Felix Ehrenhaft (24th April 1879 - 4th March 1952), who was a physician of Austria, published a different result from Millikan's in the 1911 paper, and a dispute was provoked between them (Goodstein 2001). In this dispute Ehrenhaft proposed that the elementary charge might be a fraction of Millikan's result, however Millikan refused the possibility. Today, it has been known that a quark has about one third of Millikan's result. (3) Millikan responded to Ehrenhaft's challenge, and published the 1913 paper. In this paper he omitted the data set which lowered the precision of his measurements and included the data set which raised

*) Correspondence : Institute for the Advancement of Higher Education, Hokkaido University, Sapporo 060-0817, Japan
E-mail : fumike@high.hokudai.ac.jp

its precision although the data set was out of the term which was mentioned in the paper (Franklin 1997). He didn't fabricate data set, however it seems to be a kind of scientific misconduct. And if he had taken Ehrenhaft's proposal into consideration, he might have discovered a quark.

In this paper, I showed the Millikan's psychological change from (2) to (3) based on the language information in his two papers, that is, what kinds of words and how often he used them in each paper. In a communication, a person's psychological situation is transmitted by non-verbal ways such as expressions, attitudes, colors of voice, intonations, accents and so on more than verbal way (Mehrabian 1981). However, according to Transactional Analysis (TA) which is a kind of communication psychology, it is found that there are a relation between specific words and their corresponding psychological situation (Stuart and Joins 1987). Then, I showed the Millikan's psychological change by analyzing the language information in his two papers from the viewpoint of TA. It was possible that this change conducted him to the data manipulation.

In the use of my method there are mainly three merits: the one is that it enables to detect a scientific misconduct before a misconducted paper is published, the second is that it supports an effective scientific ethical education, and the last is supporting our creativity.

It generally takes much time and money that a scientific misconduct is revealed by an exposure from within or its retests after the misconducted paper is published. By contrast, my method can reduce the time and money because it supports us to suspect a paper of doing a misconduct in the review process. In the traditional ways of reviewing papers, reviewers need not to check a misconduct but a validity of research in a paper. The author is entrusted with whether he/she does a scientific misconduct or not.

Moreover, my method supports more effective scientific ethics education than usual. In the usual education, the kinds of misconducts and these penal codes are informed to researchers in a guideline. However, even though we understand misconducts are bad, we do them in spite of ourselves when we are forced in mental. By my method we can realize such mental situation and give up to do misconducts by ourselves.

In the traditional studies on creativity psychology, the main stream is an analysis of a prominent scientist's

biographies, his/her autobiographies or both (Shun and Shunya 2003). And this method has three problems. The first is related to truth or falsehood of facts written in a biography or an autobiography (Weisberg 1986). My method can avoid this problem, because my objects of analysis are not the facts but the wordings in Millikan's scientific papers. The second is a hardship of generalizing the knowledge gained by the analysis, that is, even though we could get some standards of creativity from them, it would be difficult to prove whether the standards were valid for only prominent creative scientists or not (Wallach 1976). And the last is a difficulty of persuading all people to accept a standard of creativity (O'Quin and Besemer 1999). These two problems originate in a subjectivity of creativity. To avoid these problems I analyzed the Millikan's psychological situation when he refused the possibility of Ehrenhaft's suggestion, in other words, he disabled himself from demonstrating his creativity (Adams 1989). That is to say, I don't discuss a subjective standard of creativity, but find some negative psychological factors against creativity. Although the change of these factors doesn't guarantee us to demonstrate our creativity, it will not be negative for our creativity.

Method

First, I examined the kinds of words and each frequency in the two Millikan's papers by using CasualConc which is a free software for making an index. In this operation, the conjugations of a verb such as a present/past tense, a present/past perfect tense, a present/past participle, the present/past progressive form and so on were counted as its present tense, the plural form of a noun was counted as its singular form, and the comparative degree and the superlative degree of an adjective or an adverb were counted as its positive degree.

Next, from these words, I excluded the words which had low relations to Millikan's mental situation which led himself to manipulate data. According to TA, there are three ego-states in our minds: "Parent (P)," "Adult (A)" and "Child (C)" (pp.11-16 in Stuart and Joins 1987). "P" is a set of behaviors, thoughts and feelings copied from our parents or parent figures, and "C" is a set of them replayed from our childhoods. These two

ego-states are rooted in our past experiences, on the contrary, “A” originates in “here-and-now” situation. “A” is a set of behaviors, thoughts and feelings as our direct responses to the here-and-now as grown-ups we now are. That is, “A” has a rational function, and doesn’t relate to misconducts. So I excluded the words which corresponded to the following (a) to (p) categories, because these categories were thought to relate Millikan’s “A” to a greater or lesser degree.

- (a) A numerical recognition relates to our “A.” So I excluded numbers, numerals, quantifiers, seasonal names, month names etc.
- (b) Scientific symbols, constants, variables, units etc. also relates to our numerical recognition. So they were excluded.
- (c) In the proper use of “the” definite/“a” indefinite article, “A” seems to work consciously or unconsciously. Then they were excluded.
- (d) People’s names were used as references of the earlier researches, and they were used by Millikan’s “A” which would recognize facts in that time. So they were excluded.
- (e) Places’ names were used mainly as places where the earlier researches were done, the usage was same as (d). Then they were excluded.
- (f) Pronouns point people, things and affairs which have already appeared in a text, and they were used by our “A” which could recognize facts in that time. So, they were excluded.
- (g) Relative pronouns and interrogatives were excluded, because relative pronouns were same as pronouns and interrogatives were used to confirm facts by using “A.”
- (h) Prepositions were also excluded because their usages seemed to need “A” consciously or unconsciously.
- (i) Conjunctions express rational relations between phrases or sentences, so they were excluded because they were used by “A.”
- (j) “Be” verbs were excluded because I couldn’t judge whether they were used by Millikan’s “A” or not based on only them.
- (k) Adjectives and adverbs which signify time, place, amount etc. were excluded because they were used to confirm facts at that time.
- (l) The names of experimental equipments and materials were excluded because they were recognized by Millikan’s “A.”

- (m) The terms used in explanations of experiments depended on Millikan’s recognition of facts and phenomena at that time, so they were excluded.
- (n) To use terms on physics Millikan must work his “A,” so they were excluded.
- (o) The terms used to explain experimental results were excluded because of the same reason of (m).
- (p) The wordings which were specific for papers, such as “article,” “paper,” “summary” and so on were excluded because Millikan’s “A” worked to recognize a style of academic papers.

It was possible that the rest of words reflected Millikan’s not “A.” That is, the difference of using the rest between the two papers could express Millikan’s mental change, and the change could lead him to the data manipulation. So, lastly, I analyzed whether each frequency of the rest had a statistical significant difference or not in the two papers. If the frequency of a word is significantly higher in the 1911 paper than in the 1913 one, it means the word is peculiar to the 1911 and vice versa. To analyze the differences, I examined chi-square test on each frequency of the rest. Namely, I set an average of two frequencies of a word in each paper as its expected value, and tested the difference between the frequencies. For example, the 1911’s frequency of “exact (contains “exactly”)” was 17, and the 1913’s was 7, so its expected value became $(17+7) \div 2 = 12$. And its chi-square value became 4.17 by the calculation of $(17-12)^2 \div 12 + (7-12)^2 \div 12$. Because the test was done on the two different groups, so the degree of freedom was 1. The value of 5% probability in the chi-square distribution with 1 freedom-degree is 3.84 and that of 1% is 6.64, so there was a statistic significant difference at the level of 5% between the two frequencies of “exact,” that is, the frequency of the 1911 paper was significantly higher than that of the 1913 one. However, this way is provided in the condition that there is no statistic significant difference between the two populations. If there is a difference in a chi-square test, I have to set the average of the sizes of the two populations as their new population size, and to revise each frequency based on this new value. For example, the numbers of the rest was 1853 in the 1911 paper and 1517 in the 1913 one, so their average was 1685. Then their value of chi-square test was 33.50. Because this value meant an existence of statistic significant difference at the level of 1%, I needed to revise frequencies of each word. For instance,

in the case of “exact,” the revised value in the 1911 paper was $12 \times 1685 \div 1853 \approx 10.91$, and that in the 1913 one was $7 \times 1685 \div 1517 \approx 7.78$. Consequently, their value of chi-square revised to be 3.13, and there was no statistic significant difference even at the level of 5%.

Results

Table 1 shows the numbers of kinds of words in each category (the second column from the left), the 1911's total frequencies of words in each category (the third column from the left) and that of the 1913's (the last column from the left). The object of my analysis is the words which were contained in the rest.

As calculated in the former section, because there was a statistic significant difference at the level of 1% between the total frequencies of the rest in the 1911 paper and that in the 1913 one, I conducted chi-square test on the two frequencies of each the rest based on their revised frequencies. Table 2 shows the words and their statistic information which had significant differences. One star (*) shows that there was a significant difference at the level of 5%, and two stars (**) means there was at the level of 1%. And the background of the revised frequency which was significantly higher than the other was colored grey. The shallow grey means higher in 5%, and the dark grey shows higher in 1%.

The following characteristics of wordings in the 1911 paper were found in Table 2. The frequencies of words like “closely,” “seem,” “tend” which expressed uncertainty were higher than those of the 1913 one on the one hand and the frequencies of declarative words like “apparent,” “direct,” “account,” “assert,” “deduce,” “evident” also higher on the other. Moreover, “law” and “theory” were often referred to exclude things against them as “exceptional” things. Based on TA knowledge, we have 5 obsessions which drives us to specific says and does, that is “be perfect,” “please others,” “try hard,” “be strong” and “hurry up” (pp.155-158 in Stuart and Joins 1987). Uncertainty means vulnerability for a scientist, and Millikan cut off them by declarative words and standards. So the “be strong” driver worked on Millikan's mind mainly when he wrote the 1911 paper.

On the other hand, the characteristics of wordings in the 1913 paper were followed:

- There were many “common” and “constant” which

showed that conditions or values were unchanged.

- There were many words like “arrange,” “method,” “measurement” which were more concrete than “law” or “theory.”
- There were many objective expressions like “evaluate,” “follow,” “introduce,” “represent” which avoided declarative expressions.
- There were many expressions like “the ‘error’ or the ‘uncertainty’ was “** part of **” which also avoided declarative expressions, but insisted its accuracy.
- There were many expressions like “improve,” “reduce,” “precise” which showed improvements of correctness of elementary charge.

By these characteristics, I could conclude that the “be perfect” driver worked on Millikan strongly when he wrote the 1913 paper.

Discussion

This change of driver from the “be strong” in the 1911 paper to the “be perfect” in the 1913 one was caused by Ehrenhaft pointing out the uncertainty of the Millikan's elementary charge in the 1911 paper and the possibility that the elementary charge might be a fraction of the value of the Millikan's measurement. This possibility was very contrary to Millikan's belief on elementary charge. In the 1911 paper the “be strong” driver worked on him strongly and the driver led him to chose words that cut off ambiguities, because Millikan had a conviction for the true value of elementary charge and hurried to declare it. Then he might think he had to measure it more precisely, then the “be perfect” driver instead of the “be strong” driver led him to manipulate data after the Ehrenhaft's comment.

In this way, we found a specific Millikan's mental change to a scientific misconduct by examining the statistic difference between frequencies of words in his two papers and by analyzing the difference based on TA. By using this way, we come to be able to detect a possibility of scientific misconduct at its review process. This early detecting can reduce the social bears, such as the time and money to retest, caused by a scientific misconduct. And we can avoid falling into do misconducts by learning TA and being conscious our mental situations. Furthermore, we come to be able to find many possibilities and we can demonstrate our

Table 1. Words' categories and their statistic information

	the kinds of words	the total frequency in 1911's	the total frequency in 1913's
(a)	48	232	130
(b)	10	105	83
(c)	3	1307	1014
(d)	63	113	90
(e)	12	5	12
(f)	25	584	338
(g)	15	223	114
(h)	32	1933	1364
(i)	14	524	383
(j)	6	503	351
(k)	72	419	294
(l)	103	464	334
(m)	356	1405	998
(n)	54	386	169
(o)	84	500	418
(p)	17	58	45
the rest	381	1853	1517
Total	1295	10614	7654

Table 2. The words which had statistic significant differences

Words	Frequency in 1911's	Frequency in 1913's	Revised Frequency in 1911's	Revised Frequency in 1913's	Chi-square value
account	12	3	10.91	3.33	4.03*
apparent, apparently	7	0	6.37	0	6.37*
assert, assertion	6	0	5.46	0	5.46*
close, closely, closeness	9	0	8.18	0	8.18**
common, commonly	0	7	0	7.78	7.78**
constant, constancy	17	35	15.46	38.88	10.09**
deduce, deduction	8	0	7.27	0	7.27**
direct, directly	17	4	15.46	4.44	6.10*
error, erratic, erroneous	27	42	24.55	46.65	6.86**
evaluate, evaluation	0	8	0	8.89	8.89**
evident, evidence, evidently	11	1	10.00	1.11	7.11**
except, exceptional	8	0	7.27	0	7.27**
follow	9	21	8.18	23.33	7.28**
fundamental	0	4	0	4.44	4.44*
improve, improvement	0	6	0	6.66	6.66**
instant, instantly	5	0	4.55	0	4.55*
introduce, introduction, introductory	9	17	8.18	18.88	4.23*
law	31	11	28.19	12.22	6.31*
measure, measurement, measurably	11	31	10.00	34.43	13.43**
method	22	41	20.01	45.54	9.95**
occur	5	0	4.55	0	4.55*
opposite	5	0	4.55	0	4.55*
part	4	17	3.64	18.88	10.32**
precise, precisely, precision	5	14	4.55	15.55	6.02*
produce, product	15	3	13.64	3.33	6.26*
reduce, reduction	1	11	0.91	12.22	9.74**
represent, representation, representative	3	11	2.73	12.22	6.03*
seem	9	1	8.18	1.11	5.38*
tend, tendency	10	2	9.09	2.22	4.17*
theory, theoretical, theoretically	20	5	18.19	5.55	6.72**
uncertain, uncertainty	9	22	8.18	24.44	8.10**

creativity.

Summary

In this paper, I showed that we could read drivers of TA based on a language information in scientific papers. And we found that in case of Millikan's oil-drop experiments the change of his main driver from the "be strong" to the "be perfect" caused him to manipulate data. In the future, by examining more papers in the same way, I get more information of a characteristic mental state which cause a scientific misconduct. And based on this information, I will develop an automatic system which can judge whether a scientific misconduct contains in a paper or not.

References

- Adams, J. L. (1989), "Emotional Blocks," in Leaviitt, R.J., Pondy, L.R. and Boje, D.M.(eds), *Readings in Managerial Psychology*. Chicago: University of Chicago Press, 107-116
- Fletcher, H. (1982), "My Work with Millikan on the Oil-drop Experiment," *Physics Today*, Vol.35, Issue 6, 43-47
- Franklin, A. (1997), "Millikan's Oil-Drop Experiments," *The Chemical Educator* 2 (1), 1-14
- Goodstein, D. (2001), "In the Case of Robert Andrews Millikan," *American Scientist*, 54-60
- Millikan, R.A. (1911), "The isolation of an ion, a precision measurement of its charge, and the correction of strokes's law," *Physical Review*, Vol.32, No.4, 349-397
- Millikan, R.A. (1913), "On the elementary electrical charge and the avogadro constant," *Physical Review*, Vol.2, No.2, 109-143
- Mehrabian, A. (1981), *Silent messages: Implicit communication of emotions and attitudes*. Belmont, California: Wadsworth
- O'Quin, K. and Besemer, S.P. (1999), "Creative products," in M.A.Runco and S.R.Pritzker (Eds.), *Encyclopedia of Creativity* 1. San Diego: Academic Press, 413-422
- Stuart, I and Joins V. (1987), *TA Today, A New Introduction to Transactional Analysis*, Nottingham, England and Chapel Hill, USA: Lifespace
- Shun, Y. and Shunya, I. (2003), "Psychological Research Trend on Creativity," *NII Journal*, No.5, 65-73
- Wallach, M.A. (1976), "Tests tell us little about talent," *American Scientist*, 57-63
- Weisberg, R.W. (1986), *Creativity, genius and other myths*. New York: W.H. Freeman