

論文中の言語情報と科学的不正へ至る心理変化との関係

池田 文人†

北海道大学高等教育推進機構†

はじめに

1911年にMillikanは電気素量の測定に関する最初の論文をPhysical Reviewに投稿した[1]. この実験結果に関してMillikanとEhrenhaftとの間で論争が生じた[3]. 続く1913年の論文[2]でMillikanは測定値の精度を高めるために, 精度を下げるデータを除外し, 精度を上げる期間外のデータを入れた[4]. これは科学的不正行為につながる.

本研究は, Millikanが科学的不正行為へと至った心理状態の変化を, 1911年と1913年の論文中の言語情報の変化から読み取る. 科学的不正は, 同僚による内部告発や他の研究者による追試実験などにより明らかになる. これでは不正が行われてから不正が判明するまでに時間もお金もかかり, 社会的負担が大きい. 本研究は書かれた論文から科学的不正の可能性を検証することができるため, 社会的負担を削減できる. 従来の論文の査読では研究の妥当性は問われるものの, 不正行為の有無については, 研究者の善意に基づき問われることはない. 本研究により, 論文中の言語情報から不正行為を検出できる可能性が示されれば, 査読の段階, ある行為は研究者自身が論文を執筆する段階で, 不正行為に気づくことができ, 防ぐ事が可能となる. また, 従来の倫理指導では, ガイドラインにより, 不正行為の種類とその措置について研究者に周知される. しかし不正行為は, 頭では分かっているが, 心理的に追い込まれることにより, 不正行為になってしまうものである. 本研究では, 不正行為に追い込まれる心理状態を明らかにするため, より具体的な倫理指導が可能になる.

方法

ミリカンの電気素量を求める油滴実験に関する二つの論文[1][2]のそれぞれについて単語を切り出し, 各単語の出現回数を数える. この処理には用語索引のためのフリーソフトであるCasualConcを使用した.

次に, 交流分析に基づき, 創造性を阻害する心理的要因とは関係が低いと判断できる単語を可能な限り除外し, 分析対象となる単語を絞り込む. 交流分析によれば, 我々の心の中には, 「親 (P)」「成人 (A)」「子 (C)」の三つの自我状態がある[6]. Pとは, 我々が親もしくは親的な存在の思考・行動・感情をまねしている時の自我状態である. Cとは, 我々が子どもの時にしていたのと同じように考え, 行動し, 感じている時の自我状態である. PとCが我々の過去の記憶に基づいているのに対して, Aは「今, ここ」の現実に対して, 十分に成長した人間として考え, 行動し, 感じるような自我状態であり, 不正行為を防ぐ働きである. したがって, Aが働いていると考えられる単語を除外した.

除外された単語以外の単語について, 1911年の論文と1913年の論文にそれぞれ特徴的な単語を特定する. すなわち, 二つの論文で使用されている単語の出現頻度に統計的に有為な差があるものとないものとを分ける. 有意な差がある

単語は, いずれかの論文に特徴的な心理状態を反映している可能性がある. 出現頻度の有意差の有無を調べるために, 各単語の1911年と1913年の出現頻度について χ^2 検定を行う.

結果

除外する単語を除いた単語数は1911年が1861, 1913年が1517であり, χ^2 検定により1%水準で有意差がある. そこでこの二つの母集団の大きさの平均値1689が各集団の大きさとなるように各単語の出現頻度を補正した上で χ^2 検定を行った. その結果, 5%水準(*)あるいは1%水準(**)で有意な差があった単語は表1のとおりであった. なお, 補正した出現頻度が優位が高かった方を灰色に塗りつぶしている. 色の薄い方は5%水準で, 色の濃い方は1%水準で, それぞれ高かったことを示している.

考察

表1における1911年の論文の言語的特徴は, closelyやrate, seem, tendなど, 曖昧な意味を表す単語とともに, apparentやdirect, account, assert, deduceといった断定的表現や, lawやtheoryといった既存の権威を示す言葉, そしてそこからはずれたものをexceptionalとして切り捨てる表現が多い. このことは, 交流分析における, 人間を特定の言動に駆り立てる5つのドライバーのうち「強くあれ」が支配的に働いていると解釈できる.

一方, 1913年の論文の言語的特徴は, 条件や数値の一定性を示すcommonやconstant, evaluateやfollow, introduce, representといった断定を避ける客観的な表現, improveやreduce, preciseなど, 数値の正確さが向上したことを示す表現が多い. このことから1913年の論文作成時には「完璧であれ」というドライバーがMillikanの心の中で支配的だったと解釈できる.

このような「強くあれ」から「完璧であれ」へのドライバーの変化は, 1911年の論文における得られた数値の曖昧さをEhrenhaftにより指摘され, かつMillikanの得た数値の3分の1が素電荷であるという, Millikanの考えとは異なる可能性を指摘されたために生じた. Millikanには素電荷の値に対する確信があり, その証明を急ぐあまり, 1911年の論文では「強くあれ」というドライバーが強く働き, 多少の曖昧さを切り捨てるような言葉遣いとなった. しかし, Ehrenhaftの指摘により, 「強くあれ」だけでは自分の確信を証明することができないと分かり, 「完璧であれ」というドライバーを働かせることにより, 自分の得た数値の正確さを証明しようとした. このため, 論文に記載した期間外のデータを入れ, 数値の精度を向上させたと考えられる.

今後の課題

本研究により, 論文中の言語情報から交流分析におけるドライバーという心理状態を読み取ることができることが分かった. また, Millikanの場合, 「強くあれ」というドライバーから「完璧であれ」というドライバーに移行したことが, 期間外データの意図的の混入につながっていることも分かった. 今後は, Millikanと同じような事例について同じ方法で

“The Relation between Linguistic Information in a academic paper and the Author’s Psychological Change to Scientific Misconduct”

† Fumihito Ikeda, Institute for Advancement of Higher Education, Hokkaido University

分析を行い、科学的不正につながる行為を引き起こすのに特徴的な心理状態とその変化を明らかにする。しかし、Millikanの事例は例外的なものである。また、不正行為が発覚した論文は抹消させるため、不正行為のあった論文を分析することは難しい。そこで、当時は優れた研究と認められながら、その後間違いが確認された研究を対象に心理状態の分析を行う。

参考文献

[1] Millikan, R.A., The isolation of an ion, a precision measurement of its charge, and the correction of strokes' s law, Physical Review, Vol.32, No.4, pp. 349-397, 1911.

[2] Millikan, R.A., On the elementary electrical charge and the avogadro constant, Physical Review, Vol.2, No.2, pp.109-143, 1913.
 [4] Goodstein, D., In the Case of Robert Andrews Millikan, American Scientist: 54-60, 2001.
 [5] Franklin, A., Millikan's Oil-Drop Experiments, The Chemical Educator 2 (1): 1-14, 1997.
 [6] Stuart, I and Joins V., "TA Today, A New Introduction to Transactional Analysis", Lifespace, 1987.

表1：有意な差があった単語一覧

単語	1911年中の頻度	1913年中の頻度	1911年の補正頻度	1913年の補正頻度	χ^2 値
account	12	3	10.89	3.34	4.01*
apparent, apparently	7	0	6.35	0	6.35*
arrange, arrangement	2	7	1.80	7.90	3.84*
assert, assertion	6	0	5.45	0	5.45*
close, closely, closeness	9	0	8.17	0	8.17**
common, commonly	0	7	0	7.79	7.79**
constant, constancy	17	35	15.43	38.97	10.19**
deduce, deduction	8	0	7.26	0	7.26**
direct, directly	17	4	15.43	4.45	6.06*
error, erratic, erroneous	27	42	24.50	46.76	6.95**
evaluate, evaluation	0	8	0	8.91	8.91**
except, exceptional	8	0	7.26	0	7.26**
follow	9	21	8.17	23.38	7.34**
fundamental	0	4	0	4.45	4.45*
improve, improvement	0	6	0	6.68	6.68**
introduce, introduction, introductory	9	17	8.17	18.93	4.27*
law	31	11	28.13	12.25	6.25*
measure, measurement, measurably	11	31	9.98	34.51	13.52**
method	22	41	19.97	45.65	10.05**
occur	5	0	4.54	0	4.54*
opposite	5	0	4.54	0	4.54*
part	4	17	3.63	18.93	10.37**
precise, precisely, precision	5	14	4.54	15.59	6.07*
produce, product	15	3	13.61	3.34	6.23*
rate	13	0	11.60	0	11.60**
reduce, reduction	1	11	0.91	12.25	9.77**
represent, representation, representative	3	11	2.72	12.25	6.06*
seem	9	1	8.17	1.11	5.36*
tend, tendency	10	2	9.08	2.23	4.15*
theory, theoretical, theoretically	20	5	18.15	5.57	6.68**
uncertain, uncertainty	9	22	8.17	24.49	8.16**