

## 消えた格差——ジェンダー・バイアスが「存在すること」と「見えること」のあいだ

Rivera, L.A., Tilcsik, A. (2019) "Scaling Down Inequality: Rating Scales, Gender Bias, and the Architecture of Evaluation," *American Sociological Review* 84 (2) : 248-274.

東京大学大学院博士課程 鈴木 恭子

組織で働く多くの人々にとって、職場での業績評価はおなじみのものだ。評価の公平性は難しい課題だが、もし評価結果が評価される人の属性ゆえに偏見（バイアス）に歪められるとしたらどうだろう。さらに、評価制度の設計によって、そのバイアスの大きさが容易に変わりうるとしたら……。この論文はそうした点について、興味深いストーリーを明かしてくれる。

職場における業績評価とジェンダーの関係については多くの研究があり、女性は男性より一貫して低い評価を受けることが明らかになっている<sup>1)</sup>。たとえ彼らが同じ資格を有し、同じ成果を上げ、同じ行動をとっていたとしても、である。これを、ジェンダー・バイアスという。業績評価は昇給・昇進・解雇を左右するため、この評価のジェンダー・バイアスは、職場で男女格差が縮小しない原因のひとつと考えられてきた。このバイアスは、評価制度の設計に左右されるのだろうか——？ 著者は、評価制度を構成する要素のうちとくに「評価尺度の数」に注目し、「何段階で評価するかの違いがジェンダー・バイアスの大きさを左右するか？」と問いをたてた。この問いにこたえるため、(A) フィールド・スタディによる疑似自然実験と、(B) オンライン調査による実験を行った。

(A) では、北米のある専門職大学院における教員への授業評価が用いられる。この大学で教員評価が10段階評価から6段階評価へと変更されたイベントを外生的ショックとみなし、制度変更の前後で評価結果がどう異なるかを分析する<sup>2)</sup>。先行研究で、男性比率が高い分野でジェンダー・バイアスが大きいことが知られているため、当研究も男性優位の分野に焦点をあてる。もちいるデータには教員とコースの情報が含まれており、「同じ教員が教える同じコースへの評価結果が制度変更の前と後でどのように変化したか」を厳密に比較することができる<sup>3)</sup>。しかし、教員のスキルが

時間とともに変化する可能性もあるので、制度変更の前後における評価結果の変化が完全にジェンダー・バイアスの影響によるものとは言い切れない。

そこで (B) のオンライン調査による実験が登場する。この実験では、すべての被験者に Web 上でまったく同じ講義の台本を読んでもらったうえで、その講義の内容と教員を評価してもらう。その際配られる評価用紙は、教員の名前が男性のもの (John) と女性のもの (Julie)、それから評価の尺度が10段階のものと6段階のものとのが、被験者にランダムに割り当てられる<sup>4)</sup>。この実験において、被験者が得る講義内容の情報はまったく同一であることから、教員・コースの質は完全に一定に保たれているとみなせる。そのうえで、名前によって暗黙裡に示唆される教員の性別の違いと評価尺度の数の違いが、教員への評価にどのような影響を及ぼすかを分析することができる。

結果は興味深いもので、(A) の疑似自然実験および (B) のオンライン調査の実験とも共通していた。まず10段階評価では、諸要因をコントロールしても、女性教員への評価は男性教員よりもあきらかに低かった<sup>5)</sup>。とくに、(B) で被験者にまったく同じ台本を提示しても、教員が女性と判断された場合には男性と判断された場合よりも有意に低い評価しか得られなかった。この結果は、ジェンダー・バイアスの何たるかをシンプルに教えてくれる。ところが驚くべきことに、6段階評価に変更されたのちは、このジェンダー・バイアスは消滅してしまう。女性教員と男性教員との間で評価結果の差は見られなくなるのだ。いったいジェンダー・バイアスはどこに消えたのだろうか？

たしかに評価尺度を10段階から6段階に変更すれば、評価尺度が粗くなることで教員・コースの質の違いを反映しきれなくなるという側面もある。しかし、結果が示しているのは、そうではない。(A) で、同

じ教員・コースに対する評価結果を制度変更の前後で比較すると、女性教員にだけ制度変更後に評価の上ぶれが起きているのだ<sup>6)</sup>。(B)でも、女性教員の場合だけ10段階評価と比べて6段階評価で高得点への分布が多い。つまり10段階から6段階への変更にもなつて全体の分布がひとしく圧縮されたわけではなく、そこにジェンダー・バイアスを消滅させる「意図せざる歪み」が紛れ込んだのだ。その歪みとはなにか――。

著者はそれを、「数字に与えられた文化的な意味」と解釈する。すなわち数字は単なる数ではなく、文化的なかで特有の意味を帯びている。たとえば、欧米で10は「完全無欠」を意味するがゆえ、例外的に高く評価されないかぎり業績評価で10を得ることは難しい。実際、10段階評価において男女間に存在する格差は、じつはこの10をめぐる生じていた。(A)の結果によると、10段階評価で10と評価される確率は男性教員が女性教員の1.5倍もあったが、6段階評価で6と評価される確率は男女間で差がない<sup>7)</sup>。この違いは、最上位カテゴリにどれくらい厳しい基準を適用するかの違いから生じている。というのも、(B)で、「この教員はきわめて優秀 (brilliant) だと思うか?」という質問に対して Yes と回答したのは、10段階評価で10をつけた人の65.7%なのに対して、6段階評価で6をつけた人の場合には28.6%である。つまり、6段階評価における「6」の基準は、10段階評価における「10」よりもゆるいために、もともと男性よりも低く分布する女性教員の評価が相対的に上ぶれるのだ。しかしながら、これは人々が潜在的にいだいている女性への低い評価が改善されたことまでは意味しない。なぜなら上記の質問に Yes と答えた人の割合を全体でみると、教員が男性だと思っている場合に15.5%、教員が女性だと思っている場合に9.5%であり、この差に評価尺度による違いはみられなかったからだ。女性を男性より低く評価するジェンダー・バイアスはたしかに「存在している」のだが、6段階評価においてはそれが「見えなくなる」のだ。

かくして、10段階評価から6段階評価へという、一見すると小さな評価制度の変更が、ジェンダー・バイアスの顕在化を防ぐという結果をもたらした。控えめな変化だが、大きな意味をもつ発見でもある。私たちはこの論文からなにを学べるだろうか?

第一に、こんにちのアメリカの職場におけるジェンダー・バイアスの程度がうかがえて興味深い。人々は

いまだ女性教員を男性教員と同じくらい高くは評価できないと考えているが、「10点満点はダメでも6点満点なら」というのが現在のバイアスの温度感らしい。日本ではどうだろうか。第二に、私たちは評価制度の構築性に自覚的であるべきという点である。どんなに精緻に作られた評価制度も、ニュートラルで客観的なものではありえず、社会に埋め込まれた文化的意味づけや、評価者のもつバイアスが入り込む余地がある。私たちは評価することがもつこの不確かさとともに、その結果が現実を変えていく側面を忘れるべきではないだろう。第三に、評価することがもつ意図せざる力についてである。本論文が教えてくれたように、評価者がどのようなバイアスにさらされているかは、(事後的にみれば)設計に問題のある評価結果を分析することで明らかになった。このようにどんなに不完全な制度でも評価され記録されれば、それがいつかは分析に供されて、社会のありようを検証する手がかりとなることもあるだろう。それが、見えないものや測りたいものをあえて数量化することが持つ力、むしろ本来の意図とは無関係に備わる力なのかもしれない。

- 1) 当論文に先行研究の詳細なレビューがある。
- 2) 分析対象となるデータは、制度変更をはさむ29の期間にわたり、369教員・235コース、のべ10万5034件である。
- 3) 教員・コースの組み合わせを固定効果として扱い、観察されない異質性を統制している。ただし Fixed effect model は、これらの効果が time-invariant であることを仮定しており、ケース内で効果が変化する可能性を考慮できない。
- 4) これら  $2 \times 2 = 4$  つのパターンそれぞれについて100人ずつ、計400件の評価結果を得た。(A)のフィールド・スタディの場合と同じ分野の専門職大学院に実際に通う学生を対象とし、回答者の居住地域・大学・個人属性は広く分布している。
- 5) 男性は高得点の比率が高く、平均値も高い。このジェンダー格差は、男性優位の分野 (male-dominated field) で顕著にみられるが、それ以外では必ずしも明らかではない。
- 6) この分布の変化が、評価制度変更にもなうジェンダー・バイアスの縮小によるものではない可能性も考えられる。大学があらたに優秀な女性教員を採用した、女性教員だけが教える方を大きく改善させた、ジェンダー・バイアスが小さくなるトレンドが存在した、等々である。しかし、著者は周到な分析によって、これらの仮説を斥けている。また、この変化が制度変更とともに急激に起きたことも確認している。
- 7) 10段階評価における「10」は男性教員の31.4%に対して、女性教員は19.5%であった。6段階評価における「6」は男性教員の41.2%に対して女性教員は41.7%であった。

すずき・きょうこ 東京大学大学院学際情報学府博士課程/関西学院大学社会学部非常勤講師。主な論文に「労働市場の潜在構造と雇用形態が賃金に与える影響——Finite Mixture Modelを用いた潜在クラス分析」『日本労働研究雑誌』No.698 (2018年)。産業・労働社会学、人的資源管理論専攻。