

The Influence of Prior Beliefs on Scientific Judgments of Evidence Quality

Jonathan J. Koehler

Organizational Behavior and Human Decision Processes, 1993, 56, 23-55

★ はじめに

- ▶ 人々の信念と新しい証拠への反応について多くの研究がなされてきた
- ▶ Edwards(1968) : 規範的な方法での信念の上書き修正に失敗する
- ▶ Pitz 他 : 人々の確信は反証により揺るがない
- ▶ Lord, Ross, & Lepper(1979) : 反証に直面しても強い信念には高い耐性がある
- ▶ 反証情報の検索と想起の困難さと関係する
(e.g., Kleck & Wheaton, 1967; Koriat, Lichtenstein & Fischhoff, 1980; Ross, McFarland, & Fletcher, 1981; Snyder & Uranowitz, 1978)
- ▶ 情報の想起の困難さが要因でなくても, 反証を無視, 軽視, 信念に合うように解釈する
(Batson, 1975; Festinger, Riecken, & Schachter, 1956; Kuhn, Amsel, & O'Loughlin, 1988; Nibett & Ross, 1980; Pitz, 1969)
- ▶ Ross & Lepper(1980) : 人々は反証にセオリーバイアスな方法で反応する
 - 信念を確証する証拠を証拠として重視する
- ▶ "outcome bias"はこの考え方を支持する
 - Baron & Hershey(1988) : 医療や金融の成功と関係する決定はより高く評価される

★ THE SCIENTIST

- ▶ 科学者個人の予測や姿勢, 希望は判断や決定に影響してはならないという伝統的な考え方(Merton, 1942/1973; Scheffler, 1967)
- ▶ いくつかの臨床的な証拠がこの考え方は適切ではないこと示している
- ▶ Ian Mitroff(1983) :
 - 著名なアポロの科学者の多くは感情的に仕事に打ち込む
 - 強い信念を持っているほうが矛盾する証拠の却下を切望する
- ▶ Mahoney(1977) :
 - 科学レビューは自分の観点に対立する原稿に対して偏見を持っている
 - 科学研究の質についての判断は(特に信念が強いとき)科学者個人の信念と結果が合うかどうか依存している

★ THE MODEL

- ▶ 本研究の目的
 - Mitross と Mahoney の発見の一般化
 - 科学者の証拠の質の判断の記述的モデルのテスト
- ▶ 判断は証拠が科学者の期待に合うかどうかや、期待の強さによって中立的な基準からそれる
 - 強い信念を持つ科学者は信念を確証する研究により高い評価をつける
- ▶ 予測は以下のモデルから得られる
 - $Jq = Nq + AP$
 - ・ Jq : 研究や証拠の質の判断
 - ・ Nq : 中立的な観察者により判断された質
 - ・ A : 信念とデータの一致
 - ・ P : 信念の強さ
- ▶ モデルによる予測 (FIG. 1)
 - 弱い信念→中立の判断と少ししか変わらない
 - 強い信念→中立の判断から離れる

★ THE NORMATIVE ISSUE

- ▶ Mahoney(1977) :
 - 信念が判断に影響するのは社会科学におけるもっとも致命的で非生産的なことである
- ▶ Lord et al.(1979) :
 - 客観的な事実が知られているときや強く推論されるとき、事実が反映されている結果の研究は理論的により信用に足る
 - = 証拠の判断が信念に影響されるのは適切なことである
 - ベイズの観点から、信念は質の判断に影響する有用な情報であり、質の判断に影響すべきである
 - = 検査前オッズを尤度比に影響させる

★ Agreement Effect

- ▶ いくつかの単純化された仮説に基づく
 - 論点は 2 極化できるもの
 - 研究は片方の論点に当てはまりもう片方に当てはまらない
 - 研究方法はエラーしやすいが、あらかじめどちらかを支持するものではない
- ▶ 研究者ははじめに仮説 H を信じている(対立仮説は-H)
- ▶ 3つの命題に関連する事前の信念を持っている

- A=この研究は私の仮説や信念を支持する結果となる

- T=この研究は事実に一致した結果となる

- G=これは質のよい研究である

※ A,T,G も 2 極とする

- 確証が得られたとき, よい研究である可能性= $P(G|A)$

- 反証が得られたとき, よい研究である可能性= $P(G|\bar{A})$

▶ 関心: どのようなときに $P(G|A)$ が $P(G|\bar{A})$ より (どのくらい) 大きくなるのか?

★ Judgments about Specific Study Features

▶ 質の判断は特定の"分析的な"特徴とグローバルな"評価的な"特徴の両方についてなされる (Einhorn & Koelb, 1982)

▶ 一般的な, 評価的な判断を求められたとき, 評価を変化させる多くの原因がある

▶ 特定の特徴の評価を求められたとき, その特徴だけが損なわれている可能性は小さい

▶ 特定の特徴に関する判断がなされる場合, 科学的な質の判断における研究の結果の影響は小さくなる

★ Self-Perceptions and Normative Views

▶ 科学者の直感を調べることは興味深い

- Nasbett & Wilson(1977): 判断や行動に影響する要因に気付かない

▶ 科学者は判断が何かに影響されていることを否定すると予測される

★ GENERAL APPROACH

▶ 2つの方法を採用

1, 研究室実験

- 科学の大学院生に研究の質の評価をさせた

2, 臨床科学者の真逆の強い信念を持った 2つのグループを調査し, 質の評価をさせた

★ STUDY1

◇ Method

▶ 被験者

- 297名の科学の大学院生

- 平均 2.3年間大学院に在学

- 自然科学: 54%, 社会科学: 36%, その他の分野の科学: 10%

▶ 材料と手続き

- 20~35ページのブックレット

・ 2つの架空の科学の話題のバックグラウンド

- ・ それぞれに関する研究レポートが1つずつ
 - 架空の鳥(Stripeneck)の熱感知の器官の存在
 - K-ray と呼ばれる電磁波の存在
- 事前信念の強さ(強い, 弱い)の操作
 - ・ 事前に2ページの概要を読ませ事前信念の強さを操作
 - それぞれの存在を支持する人としらない人に分ける
 - さらに支持, 不支持の強さが強い人と弱い人を作る
 - ・ 1~100で信念の強さをチェック(50が中立)
 - 平均値 存在支持 $M_{strong} = 81.5$, $M_{weak} = 66.6$
 - 存在不支持 $M_{strong} = 16.8$, $M_{weak} = 28.4$
- 研究レポートの質(高い, 低い)と賛否(支持, 不支持)を操作
 - ・ 研究レポートの内容のいくつかの重要な点で手法を変える
 - ・ 結果と考察で片方の話題では信念を支持し, もう一方は支持しない
- 6つの分析的な判断と3つの評価的な判断をする(リッカートの7件法)
 - ・ 分析的: 手法や観察法など
 - ・ 評価的: 関連性(妥当性), 方法論的な質, 明快さ
- 判断が結果やレポートに影響されたかの質問に答える
- 統制群は要約, 結果なし
- ▶ デザイン
 - 被験者内要因…話題(Stripeneck vs. K-ray), 賛否(支持 vs. 不支持)
 - 被験者間要因…事前信念の強さ(強い vs. 弱い), 研究レポートの質(高い vs. 低い)
- ▶ 事前信念操作と分析
 - 被験者からの報告(研究レポートを読む前)から事前信念の操作が成功したか決定
 - 2つの分析
 - 1, 両方の話題で操作がうまくいった被験者のみ(賛否は被験者内要因)
 - 2, 最初の話題のみ(賛否と話題は被験者間要因)
 - ・ 以下, 分析2の結果を使用

◇ Result

- ▶ 2(賛否)×2(事前信念の強さ)×2(研究レポートの質?)の MANOVA を評価的, 分析的判断に対して実施
- ▶ 評価的判断 (FIG. 2)
 - 賛否の主効果($F(1, 280) = 10.51, p < .001$)
 - ・ 支持 > 不支持
 - 事前信念の強さと賛否の交互作用($F(1, 280) = 6.53, p < .02$)

- ・ 賛否の効果は事前信念が強いほうが大きい
- 事前信念の強さの主効果($F(1, 280) = 5.54, p < .02$)
 - ・ 強い > 弱い
 - ・ トピックに強い信念を持った人々がよりそのトピックに関する研究を欲している
 - ・ 強い事前信念を持っていて、結果に支持される被験者のみ有意に高い評価をすることによる効果
- ▶ 分析的判断
 - 賛否の主効果が境界的に有意($F(1, 262) = 2.82, p < .10$)
 - ・ 支持 > 不支持
 - 事前信念の強さと賛否の交互作用は有意に達しない
 - 計画比較ではニュートラル条件との違いを示せなかった
 - ・ 賛否の効果のなさが何に起因するかはわからない
- ▶ 2(判断の内容)×2(賛否)の被験者内 MANOVA を実施
 - 判断の内容と賛否の境界的な交互作用($F(1, 276) = 2.90, p < .09$)
 - ・ 賛否の効果は評価的質問のほうが大きい
- ▶ 自己知覚と規範的観点
 - 64%の被験者は自分が方法論的な質の評価は結果に影響されていないと感じた
 - 26%は影響された, 10%はわからない
 - 83%の被験者は科学者の方法論的な質の評価は結果によらないと感じる
 - 12%は依存が存在する, 5%はわからない
 - どの条件も同様の反応を示した
 - 賛否の効果にかかわらず反応は同じ
 - 影響されていないと信じていた被験者も他の被験者と同様に影響されている

☆ Discussion

- ▶ 結果は科学的な証拠の質の判断における賛否の効果の存在を支持する
 - 事前信念を支持する研究の質は高く評価される
 - 特に事前信念が強いとき, また評価的な判断を行うときに高く評価される
- ▶ 賛否の効果は事前信念が反証された人による厳しい判断の結果ではない
 - 主に強い信念を持ち,結果に支持される被験者により起こる
- ▶ 科学者は信念と一致する研究とそうでない研究を異なって調べる
 - 信念を支持するデータは"正しい"結果なので注意深さが減る
 - "正しい"結果が得られていない時は何か間違っているのではないかと疑いをいなく
- ▶ 被験者の自己知覚と規範的観点

- 事前信念は評価の質に影響しない・されるべきではない
- ▶ 分析的評価での賛否の効果の小ささは科学者の規範的な適切な判断の証拠

★ STUDY2

- ▶ 研究室実験では科学的な研究の質の判断は事前信念を支持するかどうか、および事前信念の強さに影響される
 - 研究室実験は人工的、被験者の知識が少ない、関心が低い
- ▶ 超能力(ESP)研究の実際の科学者の調査を行う
 - 超心理学主義者 vs. 懐疑主義者
 - 両サイドに確固たる考えを持った比較的多くの科学者がいる
- ▶ 懐疑主義者はいつも重要な証拠が不十分だと ESP を信じる人を批評している
- ▶ 超心理学主義者は懐疑主義者は早まった判断をすると反論している
 - グループと賛否の交互作用がこの議論を明らかにするかもしれない

◇ Method

- ▶ 被験者
 - 75名の超心理学主義者と 39名の懐疑主義者
- ▶ 材料
 - 調査の内容
 - ・ カバーレター：調査の性質とアウトライン
 - ・ 仮説的な超心理学研究のレポート
 - ・ レポートについての評価的,分析的判断のリッカート法の質問紙
 - ・ レポートの長所, 短所に関するコメントを求める質問紙
 - ・ 研究の結果が判断に影響したかどうかの質問紙
 - ・ 返信用の包み
 - 研究レポートは" ganzfeld"により発行された超心理学研究を手本にしてある
- ▶ 手続き
 - 6バージョンのレポートが用意された
 - ・ クオリティ(高い vs. 低い) × 結果(ポジティブ vs. ネガティブ vs. なし)
 - 超心理学主義者は 6つのうちいずれか, 懐疑主義者は結果なしを除いた 4つのうちのいずれかを送られる
 - クオリティが高いほうがレポートの内容に厳しい統制をかけた
 - 結果なしは統制群
 - レポートを読んで 6つの分析的判断と 3つの評価的判断を行う(リッカートの 7件法)
 - 評価的判断は研究室実験と同じ, 分析的判断は内容による

- レポートに関するコメントの予測
 - ・ ポジティブなコメントは自分の ESP に関する信念が支持されたレポートにおいて多い
- 結果が評価に影響したか回答

◇ Design

- ▶ 被験者間…賛否(支持 vs. 不支持), 研究の質(高 vs. 低), グループ(超心理学者 vs. 懐疑主義者)

◇ Return rates and demographics

- ▶ 返答率は超心理学者 38.5%, 懐疑主義者 29.8%(有意差なし)
- ▶ ESP への信念: 超心理主義者と懐疑主義者で両極端
- ▶ 学術的バックグラウンド
 - 超心理学者…社会科学
 - 懐疑主義者…自然科学
 - 超心理学者のほうが年季が入っている

◇ Result

- ▶ 評価的, 分析的判断において 2(グループ)×2(研究の質)×2(賛否)の被験者間要因 MANOVA
- ▶ 評価的判断 (FIG. 3)
 - 賛否の主効果($F(1, 79) = 9.20, p < .005$)
 - ・ 両グループで事前信念を支持する研究はより高い評価
 - グループ×賛否の交互作用が境界的に有意($F(1, 79) = 2.84, p < .10$)
 - ・ 懐疑主義者のほうが賛否の効果が大きい
 - 弱い事前信念の報告が少なかったため信念の強さの影響は決定できない
 - 超心理学の中立グループの評価的判断は賛否の条件の評価と有意な違いはない
 - 質の高い研究への評価は有意に高いわけではない
- ▶ 分析的評価
 - 賛否の効果は検出できず($F(1, 79) = 1.36, ns.$)
 - 各グループにおける計画対比により懐疑主義者における賛否の効果が有意 ($t(34) = 2.13, p < .04$)
 - 研究の質の主効果($F(1, 78) = 4.95, p < .03$)
- ▶ 研究レポートへのコメントの結果
 - 各研究レポートについて平均 9.2 のコード化可能なコメント
 - ポジティブとネガティブとニュートラルにコーディング

- ポジティブ、ネガティブコメントの量における 2(グループ)×2(研究の質)×2(賛否)の被験者間要因 ANOVA
 - ・ 賛否の効果なし($F(1, 58) < 1$)
 - ・ 研究の質の主効果($F(1, 58) = 10.11, p < .01$)
 - 質の高い研究ではネガティブなコメントが少ない
 - ・ グループ×研究の質の交互作用が境界的に有意($F(1, 58) = 3.40, p < .07$)
 - ネガティブなコメントの大部分は研究の質にかかわらず超心理学者によりなされた
 - 懐疑主義者は研究の質が低いときにネガティブなコメント
- ▶ 自己知覚と規範的観点
 - 65%の被験者は自分が方法論的な質の評価は結果に影響されていないと感じた
 - 21%は影響された, 14%はわからない
 - 85%は方法論的な質の評価は結果に影響されるべきではないと感じる
 - 13%は依存が存在する, 2%はわからない
 - グループ, 賛否の効果の影響はない

◇ Discussion

- ▶ 賛否
 - 科学的研究の質の判断における賛否の効果を支持する結果
 - ・ 自分たちの信念を支持する ESP の研究をより意味があり, 方法論的にしっかりして, 明快に示されていると信じる
 - 反証研究にネガティブなコメントをすることはない
 - 先行研究(Alock & Otis, 1980; Zusne & Jones, 1982; Russell & Jones, 1980)
 - ・ 科学者ではない ESP 信者は批判的でなく, 独断的で反証を思い起こさない
 - 超心理学者は懐疑主義者よりバイアスを示すという傾向はない
 - どちらかといえば懐疑主義者のほうが賛否の効果が大い
 - 分析的判断における賛否の効果の減少が見られた
 - ・ 実験の結果は特定の客観的な分析的判断に影響しない
- ▶ 自己知覚と規範的観点
 - 多くの科学者は自分の質の判断は研究の結果に影響されていない, そのような影響は望ましくないと考え
 - Nisbett & Wilson(1977)の人々は科学的な質の判断における認知過程に限られたアクセスしか持たないという結果を支持

★ GENERAL DISCUSSION

- ▶ 科学的判断における賛否の効果を支持する結果
 - 事前信念を支持するものは質が高いと判断される
 - 強い信念を持っていると賛否の効果が大きい
 - 一般的な評価的判断で顕著に見られる
- ▶ 信念への固執や現象の2極化を部分的に説明する
 - Lord et al.(1979)の死刑の研究
 - ・ 反証に晒された後は少ししか信念を変えない
 - ・ 確証にさらされたときは大げさに表現する
 - 質の判断が賛否の効果に影響されている
- ▶ The Normative Issue
 - 質の判断は可変で個人特有のものであることを暗示する
 - "事前個人内基準(Merton, 1942/1973, p. 270)"の必要性
 - ・ 訓練された科学者でも信念とその強さに影響される
 - ベイズのフレームワークの中で正当化されているがこの種の判断において適切ではない
 - ・ 賛否の効果の大きさによって多くの関数に変化する
- ▶ Direction For Future Research
 - Lord, Lepper, & Preston(1984)
 - ・ 死刑の効果性を評価するときに, 反対の結果を支持する証拠の可能性を熟考させると姿勢に一致する証拠への高評価の減少
 - Koriat et al.(1980)
 - ・ 過信は被験者の賛成する回答に反対する理由を与えることで減少する
 - 賛否の効果を減少させることは可能である
 - ・ 評価した研究に対する異なった結果をシミュレートする
 - ・ なぜ異なる結果が出たのか理由を考える
 - どのような条件で賛否の効果が起こりやすいのか?
 - 過信を減らし, 判断の質を規範的で適切な事前信念の影響の量を反映したものにするのか?
 - 証拠の質の判断の調査は科学者に限られる必要はない
 - ・ 政治, 医学, 法においても重要
 - 例) 陪審員の弁護士の論議の評価など

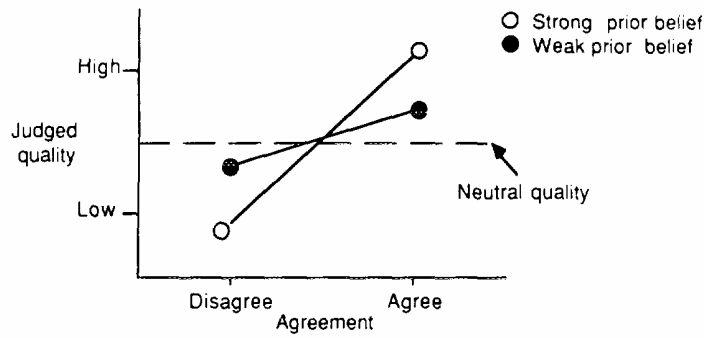


FIG. 1. Ordinal predictions for scientific judgments of evidence quality as a function of neutral quality (Nq), agreement (A), and strength of prior belief (P) as given by the model $Jq = Nq + AP$.

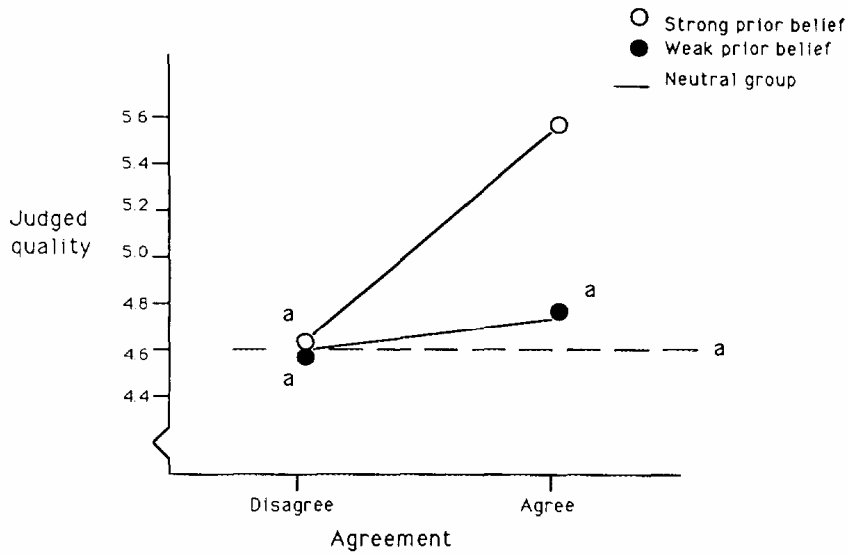


FIG. 2. Study 1 (Experiment). Mean judgments of quality for the evaluative questions as a function of agreement and prior belief strength. (a) Differs from strong prior belief/agree group, $p < .005$ (Bonferroni corrected).

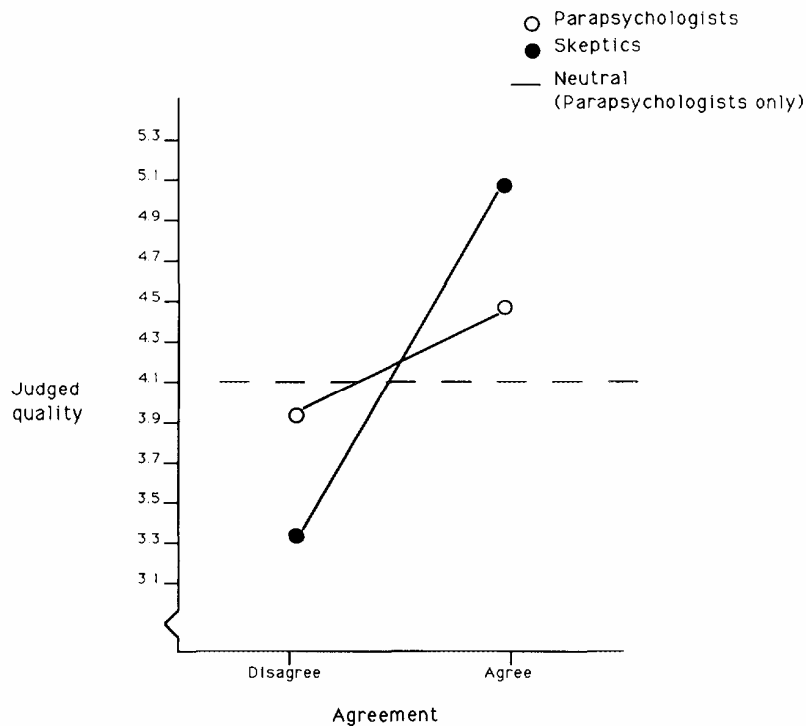


FIG. 3. Study 2 (Survey). Mean judgments of quality for the evaluative questions as a function of agreement and prior belief strength.