

Harry Potter and the sorcerer's scope:
latent scope biases in explanatory reasoning

Sangeet S. Khemlani & Abigail B. Sussman & Daniel M. Oppenheimer

Memory & Cognition (2011) Vol.39 Pp.527–535

◆ INTRODUCTION

<背景>

- 観察した行動や現象を説明する能力は、人間の合理性の特徴であり(Harman, 1965), 世界を理解し、伝達する重要な手段である(Lombrozo, 2007)。
- 近年, 説明がどのようにカテゴリー学習(category learning)や概念発達(conceptual developing), 矛盾の解消に影響を与えているかについての関心が高まっている(Ahn, Marsh, Luhmann & Lee, 2002; Murphy & Allopenna, 1994; Murphy, 2002)。

<説明の良さ(virtue)の構成因子は何か>

- 単純性(Simplicity)
 - ✧ オッカムの剃刀:
「ある事柄を説明するためには、必要以上に多くを仮定するべきでない」とする指針のこと。
- その他, virtueの構成要因であると考えられるもの
 - ✧ Lombrozo(2007)は, 人間は説明を評価する際, 「一貫性(Consistency)」 「範囲(Scope)」などを使用するかもしれないと主張。
 - ✧ これらが, 説明を生成または評価する場面において, 実際に役割を果たすかどうかを議論する研究はほとんどない。今回は, この「範囲(Scope)」について研究する。

<範囲とは>

- 別の仮説よりも多くを説明する仮説を選ぶべきである。仮説H1が2つの証拠を説明し, H2がただ1つを説明するならば, H2よりもH1を優先すべきである(Thagard, 1992)。
- ✧ より多くの観察事象を説明する理論(あるいは仮説)は, より広い(broader)説明範囲を有していて, そうでない(narrower)代替理論と比較してより強力な説明であるとされる。
- ✧ 複数の症状を併発している病人に対して, 症状ひとつひとつに病名を当てはめて説明する(症状4つに対して病名4つ)よりも, 極力少数の病名で説明する(症状4つに対して病名1つ)方が好まれる傾向がある(Lombrozo, 2007)。
- 医療診断のように, すでに観察されている現象を, 少ない説明でより多く扱っているかどうか, そしてそれが良い説明かどうかを評価することはあるが, もしかしたらあり得るかもしれない現象さえも包括する範囲の説明はどのように評価されるのかは検討されていないので, 今回実験的に調査した。
- ✧ Ex) 観測事象: ある女性が, お風呂場でネイルを塗っている。
 - 範囲の広い説明: 彼は強迫性障害である(手をお風呂場で洗うことを徹底している)。
 - 「ドアの鍵を閉めたかどうか何度も確認するかもしれない」等の非観測事象の説明もできる。
 - 説明範囲の狭い説明: 彼女はカーペットの上にネイルをこぼすことを恐れている。
 - お風呂場でネイルを塗る理由についてだけに言及している。

◆ EXPERIMENT1a-d

- 目的: 範囲の狭い説明を好むかどうかを調べるため。

<PARTICIPANTS>

- 実験1a: 30人, 1b: 31人, 1c: 30人, 1d: 26人 (年齢, 性別等の記載はなし)
- Amazon.comが運営しているオンラインプラットフォームであるMechanical Turkを通じて募集された。

<DESIGN and PROCEDURE>

- 課題遂行の動機づけのために, Harry Potter(Rowling, 1997)の物語に基づいて問題を作成した。
- 参加者には医師になりきってもらい, やってくる負傷した魔法使いたちを診断してもらう。患者らの症状(結果)は, 間違いなくDeath Eatersの呪文によって引き起こされたもので, 参加者の仕事は, どの呪文(原因)がそれらを引き起こしたのかを把握することだと伝えた。
- Delimentaという呪文は, しこり(lumps)と吹き出物(spots)を引き起こす。
Homorulaという呪文は, しこり, 吹き出物, こぶ(bumps)を引き起こす。
Darylという呪文は, しこりを引き起こすが, 斑点やこぶを引き起こすかどうかはわかっていない。
- 参加者は, 観察された効果を片方(説明範囲が広い方)だけが説明できるControl問題4題と, 両方ともが説明できるExperimental問題4題の, 計8題に取り組んだ。

Control premises	Experimental premises
A causes X. B causes X and Y. Nothing else is known to cause X or Y. Y occurred; we don't know whether or not X occurred.	A causes X. B causes X and Y. Nothing else is known to cause X or Y. X occurred; we don't know whether or not Y occurred.
A causes X. B causes X and Y. Nothing else is known to cause X or Y. X and Y occurred.	A causes X and Y. B causes X, Y, and Z. Nothing else is known to cause X, Y, or Z. X occurred; we don't know whether or not Y or Z occurred.
A causes X and Y. B causes X, Y, and Z. Nothing else is known to cause X, Y, or Z. X, Y, and Z occurred.	A causes X and Y. B causes X, Y, and Z. Nothing else is known to cause X, Y, or Z. Y occurred; we don't know whether or not X or Z occurred.
A causes X and Y. B causes X, Y, and Z. Nothing else is known to cause X, Y, or Z. Z occurred; we don't know whether or not X or Y occurred.	A causes X and Y. B causes X, Y, and Z. Nothing else is known to cause X, Y, or Z. X and Y occurred; we don't know whether or not Z occurred.

- 1a: 問題について「最も満足のいく説明はどちらか」という質問をし, その呪文名(AorB)を答えさせた。
- 1b: 「原因がA(B)である確率はどれくらいか」という尤度に関する質問をし, -3点(極めて低い)から3点(極めて低い)の7点スケールで答えさせた。
- 1c: 「患者はYをどのくらい罹患していそうか」というControl問題(Yは, A・B両方の呪文どちらともから罹患する)か, 「患者はZをどのくらい罹患していそうか」というExperimental問題(Zは説明範囲の広い呪文からのみ罹患する)のどちらかに答えた。また, 「症状が見受けられない=その症状には罹患していない」と解釈しないよう注意した。
- 1d: aと基本的に同様だが, 各呪文が必ず症状を引き起こしていたわけではなく, 確率的なものであることを考慮してもらうため, 文中の”causes”を”always causes”に変更した。また, 「患者は, AとBの呪文を, それぞれ1~5%のどれかの確率で受ける」ことを伝え, AとBの発生確率を操作した。

<RESULTS>

➤ Experiment 1a

- ◇ Control問題においては、30人中29人が、範囲の広い説明を「より満足できる説明」として選択した(二項検定, $p < .0001$)。
- ◇ Experimental問題においては、30人中25人が、範囲の狭い説明を「より満足できる説明」として選択した(二項検定, $p < .0005$)。
- ◇ 1aまとめ:
説明するために、より多くの説明が必要でないときは、範囲の狭い(説明できる事象が少ない)説明を好んで選択する。

Latent explanatory scope	Problem type	
	Control	Experimental
Broad	95	7
Narrow	5	72
Neither	0	21

➤ Experiment 1b

- ◇ 症状の原因がAまたはBであるかどうかを-3点(極めて低い)から3点(極めて高い)の7点スケールで調べた結果、問題の種類(必要な説明範囲の違い)。Control問題は説明範囲が広い説明のみが適切で、Experimental問題はどちらでも説明が可能と説明の種類(説明範囲の違い)との間に有意な相互作用があった($p < .0001$)。

Latent explanatory scope	Problem type	
	Control	Experimental
Broad	2.17	0.37
Narrow	-1.33	0.76

➤ Experiment 1c

- ◇ 「症状が見受けられない=その症状には罹患していない」と解釈しないよう操作した結果、Control問題のほうがExperimental問題より尤度測定値が有意に高く($p < .001$)、Control問題に答えた30人の参加者のうち、25人がこのパターンを示した($p < .0005$)。

➤ Experiment 1d

- ◇ 「A(B)は『常に』この症状を引き起こす」と伝え、「AとBは、それぞれ1(~5)%の確率で受ける」ことを示して呪文を受ける割合を操作した場合でも、広い説明範囲を持つ説明が選ばれた割合は、Control問題では3%、Experimental問題では80%と、結果は1aと同様のものとなった。
- ✓ 実験1a-dでは、不確実かつ不完全な情報をもとに推論する際、より広い範囲に及ぶ説明よりも、狭い範囲の説明を優先して選択することが示された。

◆ Experiment2

- 実験1dでは、不確実な要素こそあったものの、参加者が頻度情報をきちんと考慮して考えたかどうかはわからず、日常的な推論とは異なるものになってしまう。そこで、実験2では各原因の割合を操作することで、参加者が頻度情報を考慮するかどうかを検討した。

<PARTICIPANTS>

- 25人(実験1と同様の方法で集めた。年齢・性別等の記載はなし)

<DESIGN>

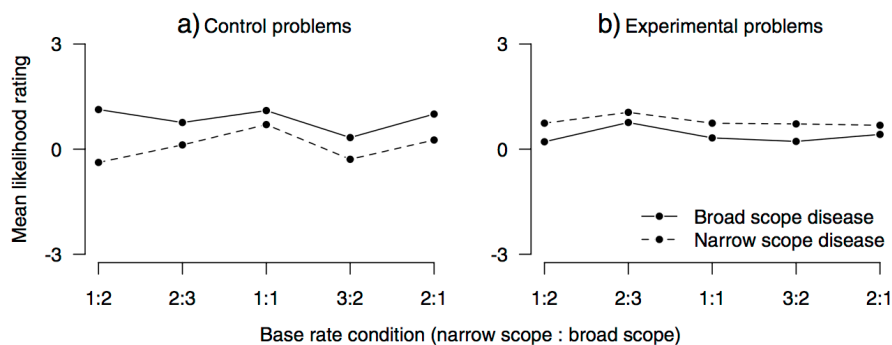
- 流れは実験1と同じ。問題のテーマこそ違う(医学)ものの、形式は実験1で用いた8題と同様のものを使用。8題の出題順序、各問題に使用された病気、患者の名前は無作為に割り当てられた。
- 参加者は、説明範囲の広い狭いの、2種類の説明が正しい可能性、+3(非常に高い)から-3(非常に低い)の7点スケールで評価した。
- 狭い説明と広い説明における頻度の比が、1:2, 2:3, 1:1, 3:2, 2:1となるように操作された。
 - ◇ Ex) あなたのもとには、1年間に約750人の患者が訪れる。そのうち、約75人の患者がVellereumと診断され、38人の患者がPythiumと診断される(ここでは、VellereumがPythiumの約2倍の頻度で発生することに注意する必要がある)。
 - ◇ 実験2では、このように2つの候補疾患の相対頻度を操作することによって、選択する説明に偏りが起こる可能性があるかどうかについて検討した。

<RESULTS>

- 実験1bでの結果と同様に、問題の種類と説明の種類との間に有意な相互作用があった($p < .05$)。
- 実験1aでの結果と同様に、Control問題においては、範囲の広い説明を(二項検定, $p < .05$)、Experimental問題においては、範囲の狭い説明を「より満足できる説明」として選択した(二項検定, $p < .005$)。

Latent explanatory scope	Problem type	
	Control	Experimental
Broad	0.84	0.40
Narrow	0.15	0.85

- 下図から、発生頻度の変化は説明範囲の広さに影響を与えないということが示された。



- しかし、日常での推論は、どの事象が何%の確率で起こり、他事象と比べてどの程度の比率であるか等はわからない。よって、実験3では、より日常的事象でも狭い範囲の説明が好まれるのかを検討する。

◆ Experiment3

<PARTICIPANTS>

- 37人(実験1と同様の方法で集めた。年齢・性別等の記載はなし)

<DESIGN>

- 参加者は、ランダムに提示される10個の予想外な事象に関する文章を読んで、それぞれに対して説明範囲の狭い説明を2つ、広い説明を2つ、計4つの候補説明のうち、どれか1つを選択した。
 - ✧ 予想外な事象は、研究者が考えたものではなく、事前に、今回の参加者とは別の参加者に「実世界で起こりうる、一風変わった出来事」を考えてもらい、また別の参加者にその出来事に対する説明をいくつか考えてもらったものを利用して用意した。
 - ✧ Ex) 出来事: ジョージは髪を黒く染めた後、全ての髪を剃ってしまった。
広い説明: 「ジョージは中年の危機(うつ病)にかかっている」等
狭い説明: 「新しい髪色が気に入らなかった」「シラミを見つけた」

<RESULTS>

- 範囲の狭い説明が選ばれた割合は64%で、広い説明が選ばれるよりも有意に頻度が高かった($p < .0001$)。
 - ✧ 提示された10の文章のうち、6つに対して当てはまり、残りの4つに対しては当てはまらなかった。そのうち、3つは説明範囲の広さで差はなく、1つは範囲が広いものが有意によく選ばれた。
- この実験3は、各文章のような出来事が起こる確率も、その選択肢が正しい確率もバラバラであるため、頑健な結果であるとは言えない。
 - ✧ しかし、実験1,2の結果を受けた上での予測と一致したため、ある程度の汎用性と確信を持って、不確実な情報下においては、範囲の狭い説明が好まれると考えられる。

◆ GENERAL DISCUSSION

- この研究でわかったことは、人間はより狭い範囲の説明を好んで選択するということである。
 - ✧ このことより、説明を生成するとき、広い範囲の説明よりも、狭い説明をより早く、より頻繁に生成する可能性が示唆される。
- 今回の結果は、先行研究の「人間はより多く説明できるものを好んで選択する」(Read and Marcus-Newhall, 1993; Preston and Epley, 2005)という主張に反するよう見えるが、それらはあくまでもすべての情報が与えられた上での選択であって、今回のように、「もしかしたら症状YやZも発症するかもしれない」といった不完全情報下での推論とは質的に異なっていると言える。
- 狭い範囲の説明でより満足するのはなぜかという疑問については、人間は不確定な情報を増やしたくないと信じていると仮定すれば、「より多く説明できる説明を選んだ場合、もしかしたらXだけでなくYも生起しているかもしれない」と考えてしまい、より不確定情報が少ない、狭い範囲の説明を好んで選ぶことが説明できる。