

Simplicity and probability in causal explanation

Tania Lombrozo

Cognitive Psychology 55 (2007) 232-257

<Introduction>

- データは説明と一貫していなければならない
 - 地動説 vs 天動説はその上で起こっている。つまり、最善の説明を選ぶためには、データとの一貫性以上のものが必要
- シャーロック・ホームズは「導き出された説明のうち、最も確率が高いものを選び」と言っているが、そもそも一般人はある事象が起こる確率を性格に判断できない
 - そのため、ある事象の説明を評価するには、「確率を生成するためのメカニズム」か「最善の説明を選択するための確率的でない規準」のどちらかが必要と考えられる
- 「単純さ (Simplicity)こそ説明を評価する規準だ」と主張する科学者もいる (ニュートン等)
 - オッカムの剃刀: 説明は、必要以上に複雑にすべきでないという考え方 (Baker, 2004)
 - 哲学では、Inference to the Best Explanation (IBE) と呼ばれる考え方 (Harman, 1965; Lipton, 2002; Peirce, 1998)
 - 「説明が複数あるときは、手持ちの証拠に最も当てはまる説明を選び」「(他が全て等しいなら,) 最も単純なものを選び」という思想
- 複雑な説明は、単純なものよりも詳細にデータを説明できるが、データにはノイズが含まれているため、新規データの予測に役立たない (一般化しにくい) 可能性がある
- 説明が比較されるときに、単純さが判断を導くかどうか、そして、そのような判断において、単純さと確率の関連については考慮されてこなかった
- ✓ Read & Marcus-Newhall (1993) では、「より少ない命題を含む説明を好む」という理論 “Theory of Explanatory Coherence (Thagard, 1989)” を実験的に検討した
 - 方法: 独立変数: 説明における要因の数 (1つ or 3つ), 目的変数: 説明の良さ
 - 手順:
 1. 参加者には、複数のデータが与えられる
(例: シェリルさんは、症状として「疲れ」「体重の増加」「腹痛」の3つを訴えている)
 2. 参加者には、単純な説明と複雑な説明が与えられ、評価してもらう
(例: 単純な説明は「妊娠」、複雑な説明は「単核球症」「運動不足」「ウイルス感染」)
 - 結果: 単純な説明のほうが、「良い説明だ」と判断された
 - 懸念: 参加者が、純粋に確率に基づいて回廊している可能性を排除することはできていない
 - ◇ たとえば、「単核球症」「運動不足」「ウイルス感染」「妊娠」の4つにそれぞれ罹患する確率が $1/4$ である場合、単純な説明 (妊娠) における確率は $1/4$ で、複雑な説明 (妊娠以外の3つに同時罹患) における確率は $(1/4)^3 = 1/64$ となり、確率的にも単純 > 複雑となる

- ✓ Lagnado (1994) では、Read & Marcus-Newhall (1993) で明らかにできなかった、説明の良さの判断に影響しているのは「単純さ」なのか「確率」なのかを検討した
 - 方法: 独立変数: ①要因の数 (疾患の数), ②確率 (各病を罹患する確率)
 - 目的変数: 説明の良さ
 - 手順:
 1. 参加者には, 「患者の A さんは, X, Y の 2 症状を訴えている. そこで担当医の B さんは, 単一の病 D1 (X と Y の両方の症状を引き起こす) か, 病 D2 と D3 の両方 (D2 は X を, D3 は Y を引き起こす) に罹患しているのではないかと考えた」という内容の文章が与えられ, どちらの説明が好ましいかに答えてもらう. その際, D1 を罹患すると, D2 と D3 を同時罹患する確率も提示される.
 - 結果: 説明の確率が提供されると, 単純さに基づいて説明を評価しない
 - 確率が $D1 > D2\&3$ のとき: D1 が好まれる
 - 確率が $D1 = D2\&3$ のとき: D1 が好まれる
 - 確率が $D1 < D2\&3$ のとき: D2&3 が好まれる
- Read & Marcus-Newhall (1993) の実験では, 確率の存在を考慮していないため, 各説明の確率についての独立した情報を提示する必要がある
- Lagnado (1994) は, 確率を考慮した上で検討を行った. しかし, Lagnado (1994) の被験者はコンピュータサイエンスの大学院生であり, その全員が統計や確率に精通していたため, このような結果になったのではないかと考えられる

<Experiment1>

- 目的: より単純な説明は好まれるか? より可能性の高い説明は好まれるか? を確かめる
- 参加者: エリート大学の学部生とサマースクール生 計 48 名 (男性 44%, 平均年齢 21 歳, $SD=3$)
- 課題: Lagnado(1994)と同内容のものを使用 (一部改変)
 - 単純さ条件 (要因の数のみで質を評価する) と確率条件 (確率の高さのみで質を評価する) がある
 - 単純さ条件 (要因の数のみが異なり, 確率に関する情報はなし)
 - 「ゾルグ星には, 750 人のエイリアンが住んでいます. あなたは, あるゾルグ星人の治療をしようとする医者です. トレダという名前のエイリアンは, 2 つの症状を訴えています. ひとつは, minttle の痛み (x) です. もうひとつは, 紫色の斑点 (y) です.
 - Tritchete 症候群 (A) は, 常に痛みを伴う minttle と, 紫色の斑点の両方を引き起こします (1,1)。
 - Morad 病 (B) は, 必ず痛みを引き起こしますが, 紫色の斑点を引き起こすことはありません (1,0)。
 - Humel 感染 (C) は, 必ず紫色の斑点を引き起こしますが, 痛みを引き起こすことはありません (0,1)。
 - その他に, minttle が痛んだり, 紫色の斑点が発生したりする原因は他には何も知られていません.
 - トレダが罹患している病として, 最も満足のいく説明は何だと思えますか?

- (A) トレダは, Tritchet 症候群である。★
- (B) トレダは, Morad 病です。
- (C) トレダは, Humel 感染です。
- (D) トレダは, Tritchet 症候群と Morad 病である。
- (E) トレダは, Tritchet 症候群と Humel 感染である。
- (F) トレダは, Morad 病と Humel 感染である。★」

- 確率条件 (確率の高さだけが異なり, 要因の数は同じ)

「ゾルグ星には, 750 人のエイリアンが住んでいます。あなたは, あるゾルグ星人の治療をしようとする医者です。トレダという名前のエイリアンは, 2 つの症状を訴えています。ひとつは, minttle の痛み (x) です。もうひとつは, 紫色の斑点 (y) です。

Tritchet 症候群は, 痛みを伴う minttle と, 紫色の斑点の両方を引き起こす。Tritchet 症候群を罹患しているゾルグ星人は, 約 50 人に存在することがわかっています。Morad 病は, 痛みを伴う minttle と, 紫色の斑点の両方を引き起こす。Morad 病に罹患しているゾルグ星人は, 約 73 人に存在することがわかっています。その他に, minttle が痛んだり, 紫色の斑点が発生したりする原因は他には何も知られていません。

トレダが罹患している病として, 最も満足のいく説明は何だと思いますか?

- (A) トレダは, Morad 病です。★
- (B) トレダは, Tritchet 症候群である。★
- (C) トレダは, Tritchet 症候群と Morad 病である。☆」

- 流れ: 参加者は, 単純さ条件と確率条件のいずれかにランダムに振り分けられた
問題冊子は 2 ページで構成されており, 1 ページ目には上記したいずれかの条件の問題文とその回答欄があり, 最も満足のいく説明を選ぶよう求めた。2 ページ目ではその回答を正当化するように求められた (なぜその選択肢を選んだかについての記述)
- 結果: 単純さ条件において, 参加者の 96% がより単純な説明を選んだ
確率条件において, 参加者の 92% がより確率の高い説明を選んだ
上記の結果は, いずれの条件においても, 選択肢のうちからランダムに選ぶよりも有意に高い確率で選ばれていると言える (いずれも $p < .01$) .
- ✓ 参加者の説明選択肢を分析するために, 正当化を 4 つのカテゴリのいずれかにコード化した
 - 単純さが直接・間接問わず示唆された場合, 単純さのカテゴリに分類
 - 単純さや, それに関連する概念を明示することなく, 単一の病気が両方の症状を説明するのに十分であると示唆した場合には, 妥当性のカテゴリに分類
 - 単純さ条件の参加者の 96% のうち, 自らの選択を正当化するために, 単純さを挙げた人が 17%, 説明に対する満足度の高さを挙げた人が 39%, 確率の高さを挙げた人が 39% だった

- しかし、後者2カテゴリのうち、そのほとんどが「トレダは2つではなく1つの疾患である」と述べていたため、単純さに気付いている可能性が考えられる
- 確率条件の参加者の92%のうち、全員が確率の高さによって自らの選択を正当化した
- 考察:
 - 単純さ情報があり、確率情報がない場合は単純さを、確率情報があり、単純さ情報がない場合は確率情報を用いて、説明の良さを判断する
 - しかし、単純さ条件における評価は、各病気における確率情報がない（各参加者において統一された基準率がない）状態での評価なので、確率情報を加えた状況で判断を下してもらうことで、単純さによって説明の良さを判断するのかどうか検討する

<Experiment2>

- 目的: 複雑説明が、単純説明よりも確率的に優位である場合、単純さと確率についてのトレードオフがどのように行われるか
- 参加者: エリート大学の学部生とサマースクール生 計 144 名（男性 40%, 平均年齢 22 歳, $SD=5$ ）
- 課題: 実験 1 で用いた単純さ条件の課題に、確率情報を組み込んだもの + 同時確率問題
 - 疾患の確率比 ($D_1 : D_2 \& D_3$) は 8 パターン
 - たとえば、一番上の $D_1 : D_2 \& D_3 = 15 : 1$ の場合は、
 $D_1 = 50/750 = 6.6\dot{6}\%$, $D_2 \& D_3 = 50/750 * 50/750 = 0.4\dot{4}\%$ となり, $6.6\dot{6} : 0.4\dot{4} = 15 : 1$ となる
 - 非単純な $D_2 \& D_3$ のほうが、単純な D_1 の確率より高くなるパターンもある

Table 1
Disease frequencies from Experiment 2

$P(D_1) : P(D_2 \text{ and } D_3)$	D_1	D_2	D_3
15:1	50	50	50
1:1	50	197	190
9:10	50	195	214
4:5	50	225	210
2:3	50	250	220
1:2	50	268	280
1:3	50	330	340
1:10	50	610	620

- 結果:
 - 単純さ vs 確率
 - ◇ 単純な説明 (D_1) を選択した参加者は 71%, 複雑な説明 (D_1 以外, $D_2 \& D_3$ 以外にも, $D_1 \& D_3$ なども含む) を選択した参加者は 29%
 - ◇ なお、複雑な説明を好んだ参加者のなかで、 $D_2 \& D_3$ 以外の説明を選択した人は、なぜその説明を選んだかを上手く正当化出来なかった（興味深いが今回はスルー）
 - ◇ 単純さが事前確率に影響を与えたと考えるモデルが最も実データを上手く表した
 - ✓ Bayesian Posterior: $P(D_1 : \text{単純な説明が選ばれる}) = P(D_1 \text{ が起こる})$

- ✓ Biased Prior: $P(D1) = P(\text{その病気が起こる}) * \alpha$ ($\alpha > 1$)
- ✓ Underweighting: $P(D1) = P(\text{その病気が起こる}) * \alpha$ ($\alpha < 1$)
- ✓ Data: 実験 2 における実測値

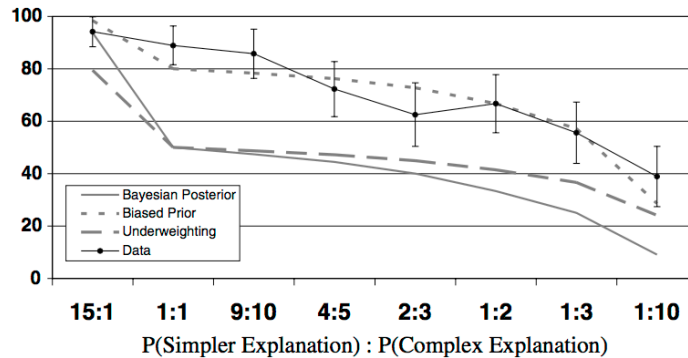


Fig. 1. Predictions and data from Experiment 2. The x-axis corresponds to the probability ratio of the simpler explanation to the more complex alternative. The y-axis corresponds to the percentage of participants selecting the simpler explanation. The data are indicated with black circles; the solid gray line presents the values that would be expected if the percentage of participants selecting the simpler explanation corresponded to the Bayesian posterior probability for the simpler explanation at the corresponding probability ratio (“Bayesian posterior”). Also illustrated are predictions corresponding to two ways in which the Bayesian calculation might be biased: by involving a prior probability favoring simpler explanations (“biased prior”) or by under-weighting the relevance of the provided probability information (“underweighting”).

- 選択した説明への正当化
 - ◇ 参加者のうち、説明を選択した理由を「単純さ」だと答えた人が 15%、「満足度の高さ」だと答えた人が 13%、「確率の高さ」だと答えた人が 51%だった
 - ◇ 9%はその他で、12%は的外れな回答をした
- 説明選択と数的能力
 - ◇ 69%が正答
 - ◇ 同時確率を正しく計算できることと、複雑な説明を選ぶことの間に関係は見られなかった ($r=.094$, $p=.263$)
 - ◇ 同時確率を正しく計算できることと、複雑な説明のほうが単純な説明よりも起こる確率が高いとき、複雑な説明を選ぶことの間にも、関係は見られなかった ($r=.15$, $p=.122$)
 - ◇ これらの結果から、同時確率の計算ができないから、確率情報を使わずに説明を評価したわけではないと言える
- 考察: 複雑説明を選択してもらうためには、4 倍以上の確率といった、偏った情報が必要
 - これは、Theory of Explanatory Coherence (Thagard, 1989)や Lagnado (1994) と異なる
 - Lagnado (1994) とは異なり、参加者には D 2& 3 の同時確率は与えられておらず、D 2&3 の同時確率が D 1 より大きいと明言されていない
 - 不確実あるいは不正確な同時確率の推定は、最初に説明を評価する方法として単純さが使用される理由の 1 つになる可能性がある (実験 4 にて検討)

<Experiment3>

- ✓ 実験 2 では、単純さと確率情報の両方が、説明の評価に影響を与えていることを示唆する結果となった。
 - 特に、単純さは事前確率に、確率情報は事後確率に影響を与えている可能性がある
 - また、単純さが確率情報を使用するキューになっている可能性もある
 - ◇ 単純な説明を選んだ参加者の半数以上が、確率の高さによってその選択を正当化したため
- ✓ 実験 3 は、基本的には実験 2 と同様だが、以下の点で異なる
 - 実験 2 では、確率情報を数値で提示したが、今回は実際にその発生頻度を体験してもらった
- 参加者: エリート大学の学部生とサマースクール生 計 108 名 (男性 44%, 平均年齢 22 歳, $SD=7$)
- 課題: パワーポイントを用いて作成. 36 パターン

“惑星ゾルグによろこそ！あなたは、ゾルグ星人の間で見つかった 3 つの病気を調べる研究チームを取り締まっている医者です。3 つの病気とは、Pilt 病, Stemmell 病, Brom 病です。”

この後、実験 2 同様、参加者には実験 2 同様、どの病気がどの症状を引き起こすかについての情報が与えられた。また、メモを取らせたため、暗記しておく必要はなかった。

“あなたの研究チームは、どの病気が稀で、どの疾患が一般的であるかを判断する必要があるのです、頻度を知ることが重要です” と述べ、各病気の発生頻度に注意を向けるよう促した。

これに続いて、各病気が陽性かどうか特定する方法の説明がなされた。各病気に対して、「額に赤いアザができる (青ではない)」「試験管から紫色の煙が出る」「黒目を光で照らすと黄色くなる」のいずれかがランダムに与えられた。

その後、参加者はスクリーンに映し出された 75 体のエイリアンを診察した。そのうちのいくつかは、先ほどレクチャーされた診断サインに従って疾患を発見することができた。

“トレダには 2 つの症状があります: 匂いのする皮膚と紫色の plickets です。トレダの症状について、最も満足のいく説明は何だと思いますか?”

参加者は、この質問に対して、6 つの選択肢のうちから 1 つを選択するよう求められた。その後、彼らは自身の選択を正当化する記述をするよう求められた。最後に、3 つの病気が何%で罹患するかを推定するよう求められた。

- 結果:
 - 単純さ vs 確率
 - ◇ 全参加者のうち、63%が単純説明を選択した
 - ◇ 単純説明の確率が小さくなるにつれて、選ばれる割合が小さくなった
 - ◇ しかし、単純説明が起こる確率 : 複雑説明が起こる確率 = 1 : 10 のときでさえ、40%以上の参加者が単純説明を選択した
 - ◇ これらの結果は、実験 2 の結果を再現したと言える

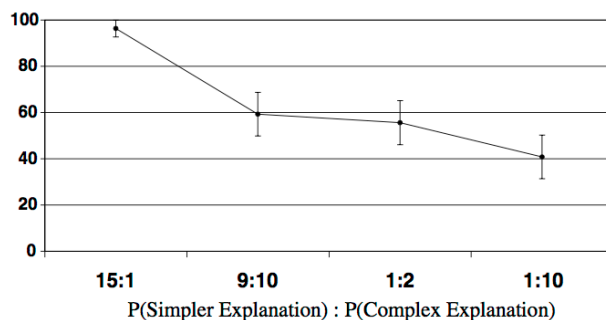


Fig. 2. Explanation choices from Experiment 3. The percentage of participants selecting the simpler explanation is indicated as a function of *base rate* condition.

- 選択した説明への正当化
 - ◇ 参加者のうち、説明を選択した理由を「単純さ」だと答えた人が 8%，満足度の高さだと答えた人が 31%，確率の高さだと答えた人が 56%，その他の理由を示した人 5%だった
 - ◇ 単純説明：複雑説明 = 15:1 あるいは 9:10 のとき、単純説明が多く選ばれた。その際、確率が正当化の理由とされた割合は 66%だった。一方、1:2 あるいは 1:10 のときは、複雑説明が多く選ばれた。その際、確率を正当化の理由とした割合は 81%だった
- 頻度情報について
 - ◇ 参加者には、数値ではなく頻度情報が提供されていたため、どの程度正しく頻度（その病気が発生する確率）を理解していたのかを調べた。また、単純説明を選んだ参加者は、その説明における病気の頻度を過大評価したかどうかについても調べた
 - ◇ Fig.3 の各パネルは、各病気（D1, D2, 3）の推定値に対応する。x 軸は単純説明と複雑説明の確率比で、y 軸は平均推定値を示す（実線が正確な値）
 - ◇ 事実、平均して、参加者は病気の頻度を正確に推定することができた
 - 答えとして D1 を選択した参加者は、15:1 と 9:10 のときは正しく頻度を推定できたが、1:2 と 1:10 のとき（D1 の確率が低いとき）は正確な値もしくは D1 以外を選んだ参加者の値と比べて、有意に高く推定されていた
 - 1:2 のとき： D1 選択者の推定値 19% vs D1 以外の推定値 8%, $t(14.6)=2.18, p < .05$
 - 1:10 のとき： D1 選択者の推定値 29% vs D1 以外の推定値 10%, $t(10.6)=1.74, p < .11$
 - ◇ これらのデータを、D1 の推定値を従属変数とし、2 * 2（1:2 or 1:10, 選択した説明が単純説明 or 複雑説明）の ANOVA を行ったところ、選択した説明についての主効果が見られた ($F(1, 50)=7.93, p < .01$)。交互作用はなかった ($F(1, 50)=.442, p = .51$)
 - したがって、単純説明を選択した参加者は、D1 の頻度を過大評価していたと言える
 - ◇ 病気の提示順序は、頻度の推定にも、説明の選択にも影響を及ぼさなかった ($p > .3$)。
- 考察
 - 実験 3 では、実験 2 よりも日常的な状況において、実験 2 の結果を再現できた
 - 数値としてではなく、経験的な形で確率情報が与えられた場合、参加者は確率情報に敏感であった（正しく確率を見積もった）が、一般的に単純説明を好んだ

- 真実であるとは考えにくい（確率の低い D1 が原因であるとする）単純説明を選んだ参加者は、複雑説明を選んだ参加者よりも、D1 の有病率に対して有意に高い推定値を示した（Fig.3 参照）
- 参加者は、説明を評価する際、単純さと確率の両方を用いたが、異なる方法で用いたと考えられる。
 - ◇ 単純さは、説明の事前確率に影響する（より単純な説明はより高い事前確率となる）
 - ◇ 事前確率に関する確実な情報が与えられていれば、事前確率に対する単純さの影響が減少するかもしれない。この可能性を実験 4 で検討する。

<Experiment4>

- 目的: 確率的に複雑説明が最も確からしいと疑う余地がない場合、単純説明に（その事前確率の）優位性を与える傾向が持続するかどうかを調べる
- 参加者: エリート大学の学部生とサマースクール生 計 144 名（男性 52%, 平均年齢 20 歳, $SD=3$ ）
- 課題: 実験 2 で用いたものとほとんど同じ
 - 検討する確率比は、単純説明 : 複雑説明 = 2 : 3 の状況のみ
 - 実験 2 では、各病気が起こる確率がそれぞれ提示されていたため、複雑説明の確率（D2&3 が起こる同時確率）を計算によって求める必要があった。今回は、そのような不確実性を排除するため、最初から「D2&3 が起こる確率」を提示
 - 参加者は、ランダムに 2 条件に割り振られた。
 - ◇ 同時確率条件: 最も満足のいく説明を選択するように求められる（実験 2 と同様）
 - ◇ 確率判断条件: 最も可能性の高い説明を選択するよう求められる
- 結果:
 - 同時確率条件では、参加者の 17%のみが単純説明を選択した
 - ◇ 同時確率条件では、参加者の 17%のみが単純説明を選択した。これは、実験 2 での対応する確率比 2 : 3 における単純説明選択率 61%より有意に低かった
 - ◇ 確率判断条件では、参加者の 46%が単純説明を選択した。同時確率条件の選択率 17%よりも有意に高かったが、実験 2 での対応する確率比 2 : 3 における選択率 61%と有意差はなし
 - 正当化
 - ◇ 単純説明を選択した参加者は、その選択の理由を「単純さ」だと答えた人が 13%、「満足度の高さ」だと答えた人が 40%、「確率の高さ」だと答えた人が 33%、その他の理由を示した人が 13%だった
 - ◇ 複雑説明を選択した参加者は、その選択の理由を「確率の高さ」だと答えた人が 94%、その他の理由を示した人が 6%だった
 - 説明選択と数的能力
 - ◇ 75%が正答
 - ◇ 実験 2 と同じで、同時確率条件において、同時確率を正しく計算できることと、複雑な説明を選ぶことの間に関連は見られなかった ($r=.05$, $p=.83$)

- ◇ しかし、確率判断条件において、同時確率を正しく計算できることと、複雑な説明のほうが単純な説明よりも起こる確率が高いとき、複雑な説明を選ぶことの間、相関が見られた ($r=.51$ 、 $p>.05$)
- ✓ 特に、単純説明を選択した参加者は、数学問題で誤る可能性が高かった
- ◇ 最も満足度の高い説明ではなく、最も可能性の高い説明を選択するように参加者に要求すると、タスクの要求が変化することを示唆している

• 考察:

- 確率情報の曖昧性を減少させた結果、多くの参加者が単純さに基づいた選択をしなくなり、複雑説明の確率が高い場合は、(これまでの実験結果と異なり)単純説明ではなく複雑説明を選択した
- 実験 2 や 3 のように、確率情報が曖昧な(提示された数値の計算が必要だったり、そもそも数値で提示されなかったりする)場合、複雑説明は同時確率を求めて評価する必要があり、評価への反映が不利になるため、単純説明によってより高い事前確率が割り当てられると考えられる
- しかし、確率情報が明白である(はじめから同時確率が数値で提示されている)場合、単純さを使用する必要性はなくなる。つまり、単純さは、確実な確率情報が存在しないときに限り事前確率に影響を与える

<General Discussion>

- 因果的な説明がいくつか存在し、その中から 1 つ選択する状況において、単純さと確率の役割を検討
- 直接的な確率情報が不在または不透明である場合に説明を評価する際、単純さに頼る (実験 2&3 vs 4)
 - 説明の単純さが上がる(より単純である)と、その説明に対する事前確率が上がる(実験 2, 3)
 - また、説明が単純であると、被説明事象が起こる確率・頻度を高く見積もる(実験 3)
 - 説明を評価することは、主観確率の推定値を生成するための仕組みとして機能するかもしれない
- 認知心理学では、因果関係の帰納についての研究が行われてきた
 - 不測の事態やメカニズムの推測のような、未知の因果関係の存在を推測する作業など(e.g. Cheng, 1997; Griffiths & Tenenbaum, 2005; Halpern & Pearl, 2000).
 - 私の知る限り、従来の研究では、単純さの役割には間接的にしか言及していないが(Lu et al., 2006; Novick & Cheng, 2004), 因果を考える際、人は原因の数に敏感であることを示唆している
- 単純さにおいて“原因の数”は心理学的にもっともらしい指標であるが、別のアプローチも存在する
 - 単純さはいくつかの方法で定量化することができ、人は他の方法に頼っている可能性がある
 - たとえば、最小記述長 (Rissanen, 1978) や Bayesian Occam's Razor (Jeffreys & Berger, 1992) のように、計算上明確な単純さの概念もある
- 今回の 4 つの実験では、参加者に候補となる説明が提示されたため、説明の評価に単純さが影響した
 - 現実的な状況では、人は説明の生成と評価の両方を行うことが多いので、単純さが評価に与える影響は小さくなる可能性がある

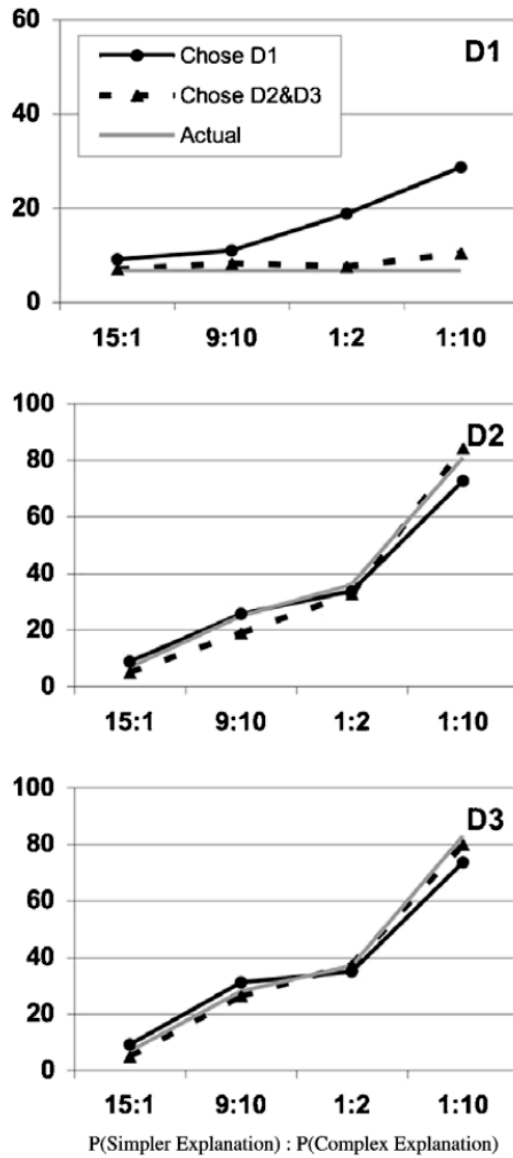


Fig. 3. Estimates of disease frequencies from Experiment 3. The average estimate for the percentage of the population with each disease is indicated. The actual values, corresponding to the data participants were shown, are indicated with solid gray lines. Participants' estimates are grouped according to whether they selected the simple, one-cause explanation or the complex, two-cause explanation as the most satisfying account of the alien's symptoms.