

# 並列単独効果が因果的説明の選好に及ぼす影響についての 実験的検討

## An Empirical investigation of Joint–Separate Effect on Preference of Causal Explanation

下條朝也<sup>1</sup> 三輪和久<sup>1</sup> 寺井仁<sup>2</sup>

Asaya SHIMOJO<sup>1</sup>, Kazuhisa MIWA<sup>1</sup>, and Jin TERAI<sup>2</sup>

<sup>1</sup>名古屋大学 <sup>2</sup>近畿大学

<sup>1</sup>Nagoya University <sup>2</sup>Kindai University

**Abstract:** What makes an explanation better than another explanation? Previous studies have suspected that explanatory virtues, such as *Simplicity* and *Scope*, affect individuals' evaluation of the explanatory goodness. Although almost all of these studies have focused on the situation that some explanations are presented simultaneously, we do not always obtain some explanations in daily life. In this research, we conducted an experiment to investigate the preference change in causal explanation between Joint and Separate Evaluations. The results showed that whether an explanation has *Latent scope* or not affects the evaluation of explanatory goodness in Joint and Separate Evaluations. Furthermore, the effect that *Simplicity* has on the judgment of goodness in explanation is different between these situations of evaluation.

### 1. はじめに

人間は説明をする生き物である[1].「説明」とは、観察された事象と、それを引き起こしたと考えられる原因の組み合わせを指す[2].我々は日常的に、身の回りで起こる様々な事象に対して、その原因を理解するために、説明を生成し、評価する。

それでは、何が説明を良いものにするのか。William of Occamは、ある事柄を説明するために必要以上の原因を仮定すべきでないことを主張し[3]、それ以後、多くの科学者が説明における *Simplicity* の重要性を主張している[4].さらに、Lombrozo (2007) は、説明受容の可否についての意思決定において、より *Simple* な説明が好んで選択されることを実験的に示した[5].

このように、人間は説明の美しさ (*explanatory virtues*) を考慮して説明の良さを考え、説明受容の可否の判断を行っていることが支持されている。 *Explanatory virtues* の要因としては、 *Simplicity* の他に、 *Consistency* や *Scope*, *Fruitfulness* が候補として挙げられる[5].このうち、特に研究されているのは *Scope* である。Thagard (1992) は、 *Scope* 以外が同じ説明同士の比較であれば、より多くの事象を説明するものを優先して選ぶべきだと主張している[6].さらに、Johnson, Toig, Keil (2014) は、事象をより多く予測する説明が好まれることを実験的に明らかにした[7].一方で、より多くの事象を予測する説明であっても、観察されない事象を予測する原因を含む説

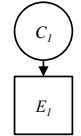
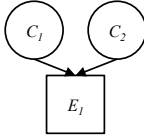
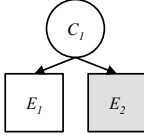
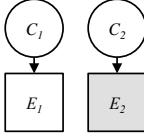
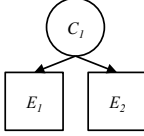
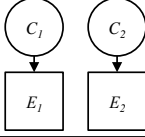
明は、含まない説明と比べて好まれないことが示唆されている[7, 8].

しかし、これらの効果は、候補となる説明が複数個存在する状況下で、それらを比較し合うことによって検討されたものがほとんどである[5, 7, 8].一方、我々は日常的に説明を生成したり、他人から与えられたりする際、必ずしも複数の説明を保持するとは限らない。それにもかかわらず、単独提示された説明の評価にどのような基準を用いるかを検討した研究は少ない。

比較可能な状況下で、説明の受容やその確率を判断する研究においては、 *Joint-Separate Effect* の影響があると考えられる[9].これは、対象を個別に提示して評価を求める単独評価 (*Separate evaluation*) と、すべての対象を同時に提示して評価を求める並列評価 (*Joint evaluation*) 間において、異なる評価が下される現象を指す。単独提示の場合、評価の際に参考となる値を持たない基準は、それがどの程度であれば「良い」と考えるべきかを判断することが困難となるため評価に影響せず、評価可能な基準のみ従って評価を下す [10].しかし、比較対象が存在する場合、評価が困難な基準であっても、その良し悪しを相対的に判断できるようになり、評価の変化が起こる。

この現象は、説明評価に関する研究においても散見される。たとえば、裁判のような状況において、容疑者 X が 1 人だけのときと、もう 1 人の容疑者 Y が存在するときでは、後者の方が、X が有罪である

表 1 本研究で用いる説明の構造

Structure	Cause(s)	Effect(s)	Unobserved Effect
	1	1	0
	2	1	0
	1	2	1
	2	2	1
	1	2	0
	2	2	0

と判断する人の割合が有意に低くなったという研究がある[11]. しかし, このとき, X や Y についての説明の質は操作されていない. そこで, Douven & Mirabile (2018, exp.1) では, Tenney et al. (2009) と同内容のシナリオを用い, 説明の質を 3 水準で操作して同様の検討を行った[12]. この研究において, 説明を単独提示した際, 質が高い説明は, そうでない説明と比べて主観的確率が高く見積もられ, 受容される割合が高かった. また, 説明を並列提示した際は, 提示された説明間の質の差が大きいほど, その効果が顕著に見られた. しかし, 説明の質はシナリオの内容の妥当性によって操作されており, explanatory virtues の要因を操作した検討はなされていない.

そこで, 本論文では, Explanatory virtues の要因として Simplicity と Scope という 2 つの要因が, 説明の良さの評価に与える影響を検討する. 特に, これら 2 つの要因の影響と, Joint-Separate Effect との関係を検討する.

## 1.1 Explanatory virtues の定義

本論文における全ての実験において, *Simplicity*, および *Scope* を, 以下のように定義する. *Simplicity* は, 説明に用いる原因の数を指す[5]. 次に, *Scope* は, 説明において原因が引き起こす事象の数を指す[7]. さらに, *Scope* は *Manifest scope* と *Latent scope* の 2 つに大別される. 前者は観測済みの事象を[13], 後者は未観測の事象を指す[8].

以下では, *Simplicity* 要因に関して, 説明に用いる原因の数が 1 つの説明を *Simple*, 原因の数が 2 つの説明を *Complex* と呼ぶ. 次に, *Scope* 要因に関して, 原因が引き起こす事象の数が 1 つの説明を *Narrow*, 事象の数が 2 つの説明を *Wide* と呼ぶ. 最後に, 未観測の事象が 0 の説明を *Manifest*, 未観測の事象が 1 つの説明を *Latent* と呼ぶ. 本実験で設定した説明の構造を, 表 1 に示す.

## 2. 実験 1A & 1B

本実験の目的は, 1 つの説明が与えられ, 他の説

明と比較できない状況（単独提示下）において、*Simplicity* 要因、*Scope* 要因が、説明の良さの評価にどのように影響するのかを検討することである。加えて、実験 1A, 1B では、事前確率を要因として操作し、上記 2 変数との交互作用を検討する。

## 2.1 方法

### 2.1.1 参加者

実験 1A には、名古屋大学に通う学部生 66 名（男性 49 名・女性 17 名、平均年齢 19.88 歳、 $SD = 0.88$ ）が、実験 1B には、同大学に通う学部生 66 名（男性 50 名・女性 16 名、平均年齢 19.59 歳、 $SD = 0.94$ ）が参加した。

### 2.1.2 実験計画

独立変数は、2 (*Simplicity*: Simple, Complex)  $\times$  2 (*Scope*: Manifest, Latent)  $\times$  3 (事前確率: 10%, 20%, 30%) で、事前確率のみ参加者間要因とした。また、従属変数として、主観的事後確率 (Subjective Posterior Probability) を用いた。また、主観的事後確率は、単独提示された説明の確率判断において用いられており [12]、説明における確率判断は、その説明の良さを測る代替の指標として用いられている [8]。

### 2.1.3 課題 & 手順

医療診断場面を扱った文章課題を作成した。具体的には、架空の症状と、その症状を引き起こしたと考えられる架空の病気の情報 (*Simplicity*, *Scope*, 事前確率) を、シナリオを通して提示し、その病気が症状を引き起こした確率 (主観的事後確率) を問うた。本実験において、各説明の事前確率は、その説明が用いる全ての原因の同時確率を提供した。

また、実験 1A では、原因が引き起こす事象の数は同じだが、観測される事象の数が異なる説明対を 2 組、計 4 つの説明 (SNM & CNM; SWL & CWL) を用いた。一方、実験 1B では、観測される事象の数は同じだが、原因が引き起こす事象の数が異なる説明対を 2 組、計 4 つの説明 (SWM & CWM; SWL & CWL) を用いた。

実験は、6 ページからなる質問紙を用いて実施された。1 ページ目で、参加者は、医師として架空の村に赴任し、訪れる各患者を診察し、それぞれが病に罹患している確率を評価することが役目であると伝えられた。その際、与えられた情報からは数理的に正確な確率を計算することはできないことも伝えられた。

その後、2~5 ページの各ページで *Simplicity*, *Scope*,

事前確率の情報が文章で提示された。以下に、例として SWM 説明 (事前確率 30%) のシナリオを示す。

診断を受けたのは、ワトソンさんです。流行っている病気のひとつは、「橙血病」という病気だと言われています。橙血病に罹患した人は、症状 E、症状 F の両方が観測されます。また、橙血病に罹患する確率はおよそ 30%だとされています。ワトソンさんを診察したところ、症状 E と症状 F の両方が見られました。

なお、*Latent scope* を含む説明では、患者から観測される 2 症状のうち、片方の症状のみ観察され、もう一方の症状は発症しているかどうかはわからなかったことが参加者に伝えられた。

各文章を読んだ後、シナリオと同ページにおいて、参加者は「このとき、X さん (対応する人名) が Y 病 (対応する病名) に罹患している確率はどの程度だと思いますか?」と問われ、患者がその病気に罹患している確率 (主観的事後確率) を回答した。この回答を *Ongoing Score* として扱った。そして、4 人の患者に対する回答 (2-5 ページ) を終えた後、全ての回答を参照して、それぞれの主観的事後確率について再度評価した (6 ページ目)。これを *Review Score* として扱った。

## 2.2 結果

### 2.2.1 記述統計

実験 1A, 1B の基礎統計量を示す (表 2 参照)。表 2 の上段の値は平均値を、括弧内の値は標準偏差を示す。後述する通り、事前確率要因は、本実験で特に検討したい 2 要因 (*Simplicity*, *Scope*) との交互作用が見られなかったため、要因として考慮せず、平均値のみを掲載した。

### 2.2.2 実験 1A における 3 要因分散分析

*Simplicity* (Simple, Complex) と *Scope* (Manifest, Latent)、事前確率 (10%, 20%, 30%) によって、単独提示された説明の主観的事後確率が異なるかを、3 要因混合計画の分散分析によって検討したところ *Ongoing Score* においては、*Simplicity* ( $F(1, 65) = 6.582, p = .013, \eta_G^2 = 0.005$ )、*Scope* ( $F(1, 65) = 89.710, p = .000, \eta_G^2 = 0.260$ )、事前確率 ( $F(2, 64) = 3.243, p = .046, \eta_G^2 = 0.060$ ) のいずれも 5%水準で主効果が有意だった。さらに、どの 2 要因間の交互作用も有意でなかった ( $p_{all} > .050$ )。

表 2 実験 1A, 1B における平均値と標準偏差

		SNM	CNM	SWL	CWL	SWM	CWM
Exp. 1A	Ongoing	63.35 (28.16)	57.73 (28.19)	33.45 (18.85)	32.26 (19.45)	-	-
	Review	64.30 (26.71)	60.77 (26.22)	35.18 (18.33)	33.95 (19.35)	-	-
Exp. 1B	Ongoing	-	-	35.15 (17.68)	31.39 (16.21)	65.08 (25.86)	61.23 (26.83)
	Review	-	-	34.98 (17.80)	31.97 (16.27)	65.38 (24.80)	62.97 (24.85)

また、Review Score においては、*Simplicity* ( $F(1, 65) = 4.968, p = .029, \eta_G^2 = 0.003$ ) と *Scope* ( $F(1, 65) = 98.494, p = .000, \eta_G^2 = 0.258$ ) において、主効果が有意だったが、事前確率 ( $F(2, 64) = 3.063, p = .054, \eta_G^2 = 0.027$ ) においては有意でなかった。交互作用はいずれも有意でなかった ( $p_{all} > .050$ )。

Cohen (1988) の提唱した基準値に基づく、両 Score とも、*Scope* が主観的事後確率の推定に与えた影響は大の基準値を上回っていた [14]。一方、*Simplicity* の効果量は小の基準値を下回っており、かつその検定力が 1 に近いため (Ongoing:  $power = 0.951$ ; Review:  $power = 0.950$ )、標本数の影響が強く見られた可能性があり、 $p$  値の解釈を行う意義が薄いと判断した。

### 2.2.3 実験 1B における 3 要因分散分析

Ongoing Score においては、*Simplicity* ( $F(1, 65) = 4.968, p = .000, \eta_G^2 = 0.008$ )、*Scope* ( $F(1, 65) = 98.494, p = .000, \eta_G^2 = 0.340$ )、事前確率 ( $F(2, 64) = 5.942, p = 0.004, \eta_G^2 = 0.116$ ) のいずれも主効果が有意だった。どの 2 要因間の交互作用も有意でなかった ( $p_{all} > .050$ )。

また、Review Score においても、*Simplicity* ( $F(1, 65) = 7.857, p = .007, \eta_G^2 = 0.005$ )、*Scope* ( $F(1, 65) = 206.564, p = .000, \eta_G^2 = 0.366$ )、事前確率 ( $F(2, 64) = 3.964, p = .024, \eta_G^2 = 0.087$ ) のいずれも主効果が有意だった。どの 2 要因間の交互作用も有意でなかった ( $p_{all} > .050$ )。

Cohen (1988) の目安に従うと、両 Score とも、*Scope* が主観的事後確率の推定に与えた効果は大の基準値を上回っていた。一方、*Simplicity* の効果量は小の基準値を下回っており、かつその検定力が 1 に近いため (Ongoing:  $power = 0.951$ ; Review:  $power = 0.954$ )、 $p$  値の解釈を行う意義が薄いと判断した。

## 2.3 考察

実験 1A, 1B の両 Score において、ともに *Scope* の影響が強く見られた。このことから、*Latent scope* を含まない説明は、含む説明と比べて、より受容される確率が高いと言える。

一方、*Simplicity* の影響は、実験 1A, 1B の両 Score において極めて限定的なものだった。この結果は、説明が並列提示された状況を扱った Lombrozo (2007) の結果とは異なるものだった。本実験で、*Simplicity* 要因の効果が極めて限定的であった理由としては、説明が単独提示されたことが考えられる。そこで、実験 2 では、並列提示下における *Simplicity* の影響を検討する。

## 3. 実験 2

本実験の目的は、説明の良さの評価における Joint-Separate Effect を検討することである。そのために、説明を並列提示し、*Simplicity* あるいは *Scope* の異なる候補説明と比較できる状況において、どちらの説明がより高い割合で受容されるかを調べた。比較の結果、*Simplicity* の良し悪しを相対的に判断できるようになり、その影響が前実験よりも強く見られるようになり、先行研究 (Lombrozo, 2007) と同様、より Simple な説明が好まれるようになると予測した。

### 3.1 方法

#### 3.1.1 参加者

名古屋大学に通う学部生 30 名 (男性 14 名、女性 16 名、平均年齢 18.80 歳、 $SD = 0.79$ ) が参加した。

#### 3.1.2 実験計画 & 課題 & 手順

本実験は、*Simplicity* が高い、あるいは *Latent scope* を含まない説明を、そうでない説明よりも良い説明であると仮定し、より多くの参加者がそれらを選択

表 3 実験 2 で用いた 7 組の説明ペア

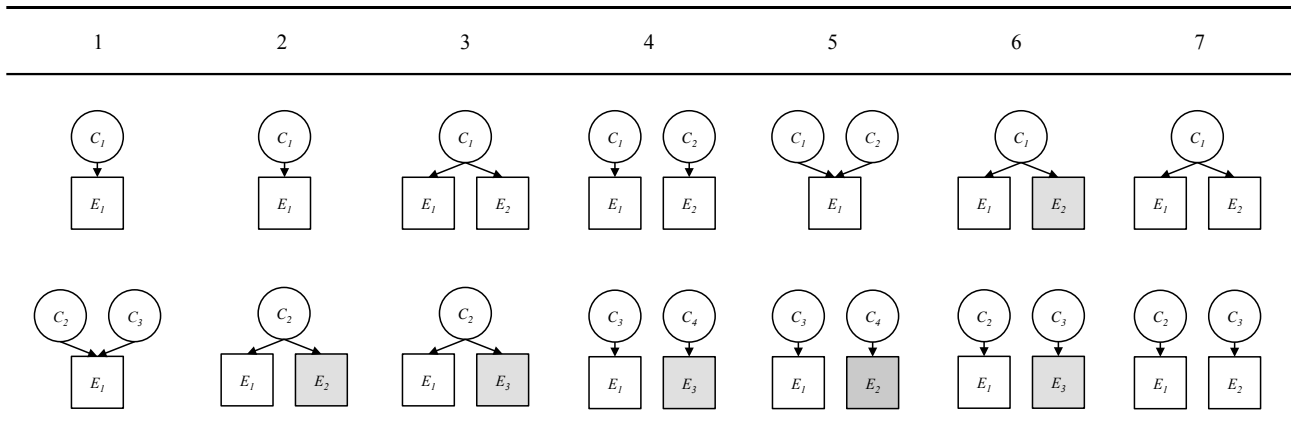


表 4 表 3 における上段の説明が選択された数と、その統計値

	1	2	3	4	5	6	7
Exp. 3	21 $p = .021$ $g = 0.20$	25 $p = .000$ $g = 0.33$	29 $p = .000$ $g = 0.47$	29 $p = .000$ $g = 0.47$	21 $p = .021$ $g = 0.17$	20 $p = .049$ $g = 0.17$	25 $p = .000$ $g = 0.33$

上段の値は、表 3 でより良いと仮定した説明が選択された数を示す。中・下段は、Fisher's exact test を実施した結果である ( $g$  は効果量)。

するという仮説のもと実施した。

実験は、8 ページからなる質問紙を用いて実施された。1 ページ目で、参加者は、架空の村に赴任した医師として、患者が、2 つの病気のうち、どちらの病気に罹患しているかを判断することが役目であることが伝えられた。

その後、2~7 ページの各ページにおいて、以下 7 つの説明のペアが 1 組ずつ提示され (表 3 参照)、参加者は、患者がどちらの病気に罹患しているかを判断した。なお、事前確率は全ての説明において等しい値 (20%) が割り振られた。原因の数が異なる説明同士を比較したペア 1, 6, 7 では *Simplicity* 要因の効果を、引き起こされると予測される事象あるいは未観測の事象の数が異なる説明同士を比較したペア 2, 3, 4, 5 では *Scope* 要因の効果を検討した。ペア 2 は Johnston et al. (2016, exp.3), ペア 7 は Lombrozo (2007, exp.1) で、同様の構造が検討されている。ペア 3 は Johnson et al. (2014, exp.1B) と、ペア 4 は Khemlani et al. (2011, exp.1a-d) と同構造だが、他の Structures との整合性を図るために、*Scope* を 3 つから 2 つに変更して採用した。一方、ペア 1, 5, 6 は先行研究で検討されていないが、ペア 1 と 6 は *Simplicity* の高さによって、ペア 5 は *Latent scope* の有無によって、上段が選ばれると予測した。

### 3.2 結果

Fisher's exact test (one-sided) の結果、全ての説明ペアにおいて、*Simplicity* が高い、あるいは *Latent scope* を含まない説明のほうが、そうでない説明と比べて、有意に多く選択された (表 4)。また、ペア 2, 3, 4, 7 は、それぞれに対応する前述の先行研究と同様のパターンを示した。ただし、効果量を見る限り、前実験と同様、*Simplicity* の影響よりも、*Latent scope* の有無による影響のほうが強かった。また、より Simple な説明が、そうでない説明と比べてより多く選択されたことから、本実験における *Simplicity* の影響は、前実験のそれよりも高まったと考えられる。また、*Latent scope* の影響は、比較の有無を問わず、依然として高かった。

## 4. 総合考察

本論文では、*Simplicity* と *Scope* という 2 つの要因が、説明の良さの評価に与える影響を検討した。

全実験を通して、*Latent scope* の有無による影響が強く見られた。実験 1B において、*Latent scope* を含む説明は、含まない説明よりも得られる evidence の量 (effect の数) が少ないため、その説明の良さが低く見積もられた可能性がある。一方で、実験 1A においては、*Latent scope* を含む説明と含まない説明がそれぞれ得る evidence の量は等しいにもかかわらず、

前者は後者と比べて主観的事後確率を低く推定された。そのため、evidenceの量ではなく、Latent scopeを含むかどうか推定に影響を与えた可能性が示唆される。一方、Simplicityの影響は、単独提示下(実験 1A, 1B)ではほとんど見られなかったが、並列提示下(実験 2)では、より Simple な説明が好んで選択された。これは、Simplicity が単独では評価困難な基準であり、他の対象における同基準と比較することによって、相対的に判断することが可能になった結果であると考えられる。

このことから、以下の 2 点が明らかとなった。ひとつは、単独・並列を問わず、説明受容の基準として高い影響力を持つのは、Latent scope であるということである。もうひとつは、説明は、単独で評価した場合と、並列で評価した場合で、用いる基準が変化するということである。また、実験 1A, 1B において、Simplicity, Scope, 事前確率のいずれの間における交互作用も見られなかったことから、Simplicity と Scope, 事前確率は、それぞれ独立して判断に影響を与えていることが示された。

さらに、説明を並列提示すること自体は従来行われてきた実験であるが、各研究における Scope の統一性はなく、説明の各ペアの効果量の比較などを行うことが難しい状態であった。そのため、先行研究で扱われてこなかった組み合わせを含め、Simplicity と Scope を要因として操作することで網羅的に検討し、比較可能な形で報告したことも、本論文における新規性の一つである。

最後に、実験 1A, 1B における Review Score と実験 2 の比較について述べる。実験 1A, 1B では、単独提示下における説明の良さの評価を検討したが、それに対応する測定変数は Ongoing Score であり、Review Score は実質的に並列提示の状況に対応している。ただし、実験 2 における比較との相違点として、Review Score では、全ての説明の評価が終わったのちに、振り返りとして並列評価をさせる状況を扱い、実験 2 では、ある説明の評価を、他の説明と比較しながら行う状況を扱った。その結果、比較の効果は、後者においてのみ観察された。このことから、Joint-Separate Effect は、評価を行うその時でないと現れないことが示唆された。

## 謝辞

本研究は JSPS 科研費 15H02717 の助成を受けたものです。

## 参考文献

[1] Norman, D. A. (1988). *The Design of Everyday Things*, Basic Books.

- [2] Salmon, W., (1984). *Scientific Explanation and the Causal Structure of the World*, Princeton: Princeton University Press.
- [3] Jeffreys, W. H. and Berger, J. O. (1992). Ockham's razor and Bayesian analysis. *American Scientist*, 80(1), 64–72.
- [4] Peirce, C. S. (1998). *The essential Peirce: Selected philosophical writings, 1893–1913*. Bloomington, IN: Indiana University Press.
- [5] Lombrozo, T. (2007). Simplicity and probability in causal explanation. *Cognitive Psychology*, 55, 232–257.
- [6] Thagard, P. (1992). *Conceptual revolutions*. Princeton: Princeton University Press.
- [7] Johnson, S. G. B., Johnston, A. M., Toig, A. E., & Keil, F. C. (2014). Explanatory scope informs causal strength inferences. *Proceedings of the 36th Annual Conference of the Cognitive Science Society*, pp. 2453–2458.
- [8] Khemlani, S. S., Sussman, A. B., & Oppenheimer, D. M. (2011). Harry Potter and the sorcerer's scope: latent scope biases in explanatory reasoning. *Memory & cognition*, 39(3), 527–535.
- [9] Hsee, C. K., Loewenstein, G., Blount, S., & Bazerman, M. (1999). Preference Reversals Between Joint and Separate Evaluations of Options: A Review and Theoretical Analysis. *Computers in Human Behavior*, 23(4), 1809–1819.
- [10] Bohnet, I., van Geen, A., & Bazerman, M. (2015). When Performance Trumps Gender Bias: Joint vs. Separate Evaluation. *Management Science*, 62(5), 1225–1234.
- [11] Tenney, E. R., Cleary, H. M. D., & Spellman, B. A. (2009). Unpacking the doubt in “beyond a reasonable doubt”: Plausible alternative stories increase not guilty verdicts. *Basic and Applied Social Psychology*, 31(1), 1–8.
- [12] Douven, I., & Mirabile, P. (2018). Best, second-best, and good-enough explanations: How they matter to reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(11), 1792–1813.
- [13] Read, S.J., & Marcus-Newhall, A. (1993). Explanatory coherence in social explanations: A parallel distributed processing account. *Journal of Personality and Social Psychology*, 65, 429–447.
- [14] Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.