

容姿の自己評価は他者からの評価と一致するのか —自計式調査による測定の可能性—

小 林 盾

[要旨]

この論文では、容姿（ルックス、美容、身体的魅力）の自己評価が他者からの評価（他者評価）と一致するのかを検討する。これまで、容姿は他計式調査で調査員が他者評価することが多かった。しかし、自計式調査で本人が容姿を自己評価できるようになれば、データ収集の範囲がおおきく広がる。そこで、20名の被評価者（20代）が容姿を自己評価したあと、無関係な16名の評価者（20代～60代）が被評価者1人1人について他者評価を与え、平均を求めた。その結果、（1）分布より、容姿の自己評価と平均他者評価に平均の違いはなかった。（2）他者評価の平均の信頼区間より、被評価者のうち半分ほど（20名中8名）は、95%信頼区間に入っていた。（3）平均他者評価を従属変数とした回帰分析より、平均他者評価を予測するには自己評価のみを用いればよく、自己評価が高い人ほど平均他者評価が高かった。以上から、容姿の自己評価は他者評価とおおむね一致し、人びとは自分の容姿をおおむね正確に把握していた。したがって、容姿の測定には、自計式調査でもある程度信頼でき、データ収集の可能性があることが分かった。

[キーワード]

容姿、評価、自計式調査

1 目的

1.1 先行研究

容姿（ルックス、美容、身体的魅力）は、どのように測定できるのだろうか。Hamermesh (2011) は、北米社会では容姿がよいと職業達成、収入達成、家族形成が促進されることを示した。容姿がよい人ほど収入が高く、容姿がよいことによるご褒美（プレミアム）よりよくないことによる罰（ペナルティ）のほうが大きい。こうした効果は男性に強くみられる（Hamermesh and Biddle 1994）。容姿のよい女性は教育の高い男性と結婚し、夫婦の容姿には正の相関がある（Buss 1994）。容姿がよい人ほど、満足して幸福である（Umberson and Hughes 1987）。

ただし、これらで容姿は、面接調査という「他計式調査」で調査員が評価していた（この論文では、こうした他者からの評価を「他者評価」とよぶ）。そのため、この形だと留置調査、郵送調査、インターネット調査などの「自計式調査」では容姿を測定できず、データ収集の範囲が制約されてしまう。

これにたいし、小林・谷本（2016）は日本人を対象にして、郵送調査という「自計式調査」で、本人が容姿を評価する形でデータ収集した（この論文ではこうした本人による評価を「自己評価」とよぶ）。その結果、容姿がよい人ほど、役職につきやすく、所得が増え（所得は男性のみ）、多くの人から告白され、交際人数や結婚のチャンスや子どもの数が増えた。さらに、高い階層にいると感じ、自信があり、幸福だった（幸福は男性のみ）。

たしかに、自己評価であれば幅広くデータ収集できる。しかし、容姿の評価とは本来的に、本人ではなく他者がキャスティング・ボート（決定権）を持つものである。そのため、自己評価が他者評価と一致している保証はない。いくら本人が「自分はハンサムだ」とか「自分はちっとも美人ではない」と思っている、周囲の人もそう感じていなくてはならない。行動や経験や心理についての質問は、おおむね回答者本人がキャスティング・ボートを持つので、この違いが容姿の測定の特徴といえる。

なお、人びとの社会生活における容姿の役割としてこれまで、女性ほど容姿へのプレッシャーが強く（Wolf 1991）、女性ほど美容整形に関心があり（谷本 2008）、男女ともに恋愛や結婚に容姿が重要である（山田・白河 2008）ことが指摘されてきた。Bjerk（2009）は理論的に、結婚市場で女性が自分の容姿を男性の所得と交換することを示した。

1.2 リサーチ・クエスチョン

そこで、この論文では以下のリサーチ・クエスチョンにアタックする。もし容姿の自己評価という主観的评价が客観性を持ち、信頼できることが検証されれば、自計式調査でも容姿を測定できるようになる。しかし、もし未解明のままだと、ともすれば容姿のデータ収集が他計式調査に限定されるため、データの蓄積が制約されてしまうだろう。

リサーチ・クエスチョン。容姿を本人が評価した「自己評価」は、同じ人を他者が評価した「他者評価」と一致するのか。

では、これまで容姿はどのように測定されてきたのか。Hamermesh（2011）の整理によれば、多くの研究で全身や服装より「顔」が容姿をよく表すとされてきた。その際、5段階または10段階で他者が評価してきた。出発点はCampbell et al.（1971）のQuality of American Life調査における5段階評価である。選択肢は「とてもハンサムまたは美人」「年齢と性別で平均以上のよい容姿」「年齢と性別で平均的容姿」「年齢と性別で平均以下のきわめて地味」「不器量」だった。

さまざまな分析の結果、容姿の評価はおおむね評価者のあいだで一致した。さらに、同じ人の容

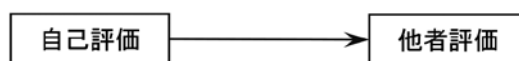


図 1 仮説

(注) 矢印は因果関係を表す。



図 2 自己評価のデータ収集の様子 (左), 他者評価のデータ収集の様子 (右)

(注) プライバシー保護のため、左の写真はイメージ。右の写真はブリーフィング時。

姿を、異なる写真や異なる時点で評価しても、似た評価となった (Hamermesh 2006)。分布では、「容姿がよい人」のほうが「そうでない人」より多かった。男女で平均に差はないが、女性の評価ほうが極端で散らばりが大きい傾向がある (Hatfield and Sprecher 1986)。また、若い人のほうが容姿がよいと評価される傾向がある (Etcoff 1999)。

1.3 仮説

リサーチ・クエスチョンから、どのような仮説を立てるべきだろうか。自己評価が他者評価と一致しているかどうかを解明するために、以下の仮説を検証する (図 1)。

仮説。容姿の自己評価が高い人ほど、他者評価が高いただろう。

分析の結果、もし自己評価が高くても低くても他者評価に違いがなかったり、自己評価の高い人ほどかえって他者評価が低かったりしたら、この仮説は支持されない。

2 方法

2.1 自己評価のデータ収集

自己評価と他者評価を、どうデータ収集できるか。この論文では、20名の「被評価者」が容姿を自己評価したあと、無関係な 16名の「評価者」が被評価者 1人 1人について他者評価を与えた。

まず、自己評価データを 2016年 1月に、20代の男女 20名 (うち男性 5名、女性 15名) から収

集した。集合調査とし、互いに回答が見えないよう十分に間隔をとって着席し、A4判1ページの調査票を用いて自計式で記入した(図2左)。リード文として「自己評価と他者評価の関連について調べています。正しく分析するために、絶対に遠慮しないで回答してください。いつでも回答を中止できます。入力後すぐにシュレッダー廃棄しますので、『だれがどう回答したか』は絶対に分かりません」とした。

小林・谷本(2016)と同様に、選択肢を下=1から上=10の10段階として、自己評価を以下の質問で測定した。

質問1(容姿の自己評価)。かりに現在の日本社会が、ルックス(顔)で上から下まで次のようなグループに分かれるとすれば、人びとはあなたをどれに入れそうですか(○は1つ)

下					上				
1	2	3	4	5	6	7	8	9	10

同時に、以下のように「一番低い回答」と「一番高い回答」が同じ1~10のどれになると思うかを質問した。この論文では「最低推測」、「最高推測」とよぶ。

質問2(容姿の自己評価の最低推測、最高推測)。実際に他の人にあなたの写真を評価してもらった場合、「一番低い回答」「一番高い回答」はどれになると思いますか(○はそれぞれ1つ)

	下					上				
一番低い	1	2	3	4	5	6	7	8	9	10
一番高い	1	2	3	4	5	6	7	8	9	10



2.2 他者評価のデータ収集

つぎに、この20名について、他者評価データを2016年12月にある団体の全職員16名(うち男性4名、女性12名、30代1名、40代8名、50代4名、60代3名)から収集した。留置調査とし、A4判カラー4ページの調査票を用いて自計式で記入した(図2右)。リード文は質問1と同じものを使用した。

そのうえで、以下の質問を用いて測定した。被評価者(自己評価者)20名の写真が並び、1人1人について10段階で評価することが求められる。20名は男女混合で、ファースト・ネームのアルファベット順にアからトまで配置されている。写真は、被評価者から就職活動で使用した証明写真を提供してもらうことで、条件を整えた(以下ではプライバシー保護のため写真をイラストへ変更した)。

質問3(容姿の他者評価)。かりに現在の日本社会が、ルックス(顔)で上から下まで次のよう

なグループに分かれるとすれば、人びとは以下の方をどれに入れそうですか (○はそれぞれ1つ)。可能なかぎり、それぞれ評価を変えてください。アからトまで20人いますので、全員を見てから記入してください。

		下										上
ア		1	2	3	4	5	6	7	8	9	10	
イ		1	2	3	4	5	6	7	8	9	10	
⋮	⋮	1	2	3	4	5	6	7	8	9	10	

2.3 分析方法

20名の被評価者の1人1人に、16名の評価者による他者評価がある。そうした他者評価の平均を求め、それをある被評価者の「平均他者評価」とよぶ。分析では、仮説に基づき、平均他者評価を従属変数とする ($N = 20$)。

独立変数は自己評価とする。回帰分析を行い、自己評価が平均他者評価をどれだけ予測できるかを分析する。統制変数として、被評価者の性別 (男性ダミー、男性 = 1、女性 = 0) と、最低推測、最高推測を用いる。頑健性のチェックでは、他者評価の中央値も用いる。

3 分析結果

3.1 分布と散布図

従属変数である平均他者評価と、独立変数である自己評価の分布は、図3左となった ($N = 20$)。記述統計は表1となった。平均他者評価は、平均5.4、中央値5.2、標準偏差1.6、最小2.6、最大8.0だった。自己評価は、平均5.5、中央値6.0、標準偏差1.3、最小3.0、最大7.0だった。対応のある平均の差の検定を行ったところ、平均他者評価と自己評価に差があるとはいえなかった ($N = 20$ 、 $p = 0.902$)。

自己評価と最低推測の相関係数は0.582 (以下すべて $N = 20$ 、 $p = 0.007$)、自己評価と最高推測では0.662 ($p = 0.001$)、最低推測と最高推測では0.831 ($p = 0.000$) だった。

平均他者評価と自己評価の散布図は、図3右となった。単回帰式を求めたら、平均他者評価 =

表 1 記述統計

		平均	中央値	標準偏差	最小値	最大値
他者評価	平均他者評価	5.4	5.2	1.6	2.6	8.0
	他者評価の中央値	5.4	5.5	1.7	2.0	8.0
	他者評価の最小値	2.7	2.0	1.4	1.0	5.0
	他者評価の最大値	8.5	9.0	1.5	5.0	10.0
	他者評価の標準偏差	1.6	1.7	0.3	1.0	2.1
自己評価	自己評価	5.5	6.0	1.3	3.0	7.0
	最低推測	3.2	3.0	1.4	1.0	6.0
	最高推測	7.4	7.5	1.4	5.0	10.0

(注) $N = 20$ 。他者評価の5変数は、評価者 $N = 16$ が被評価者 $N = 20$ に評価を行い、平均、中央値、最小値、最大値、標準偏差を求めた。

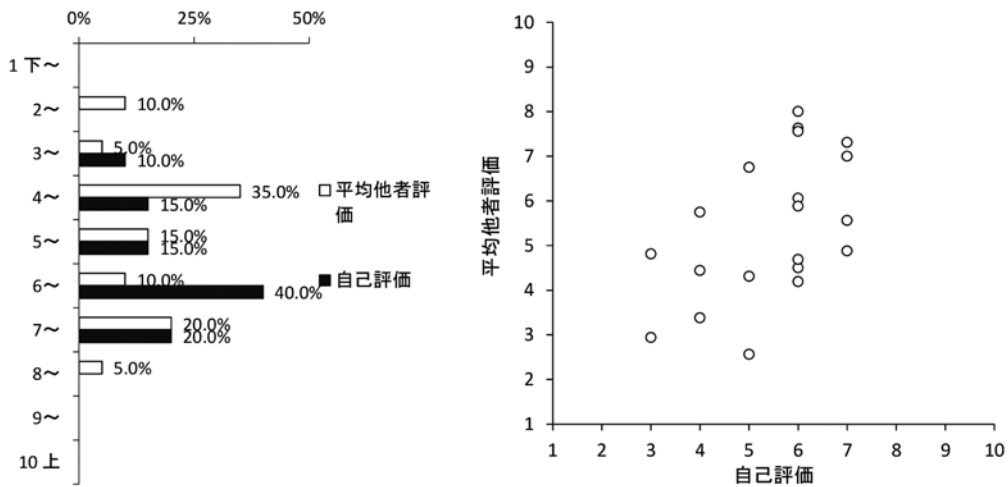


図 3 平均他者評価と自己評価の分布 (左)、自己評価と平均他者評価の散布図 (右)

(注) $N = 20$ 。

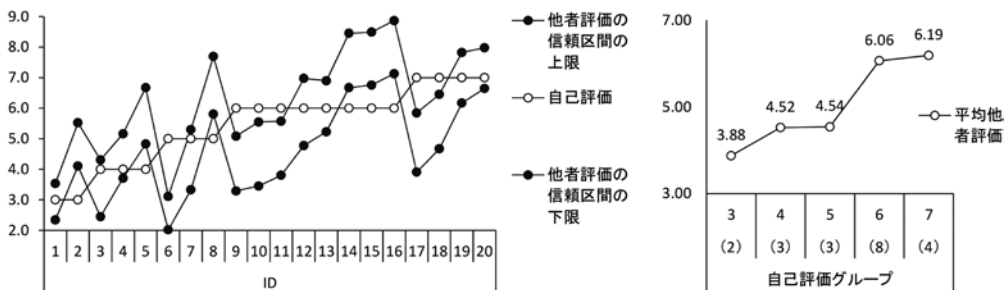


図 4 被評価者ごとの他者評価の平均の信頼区間 (左)、自己評価グループ別の平均他者評価の平均 (右)

(注) $N = 20$ 。() 内は人数。ID は自己評価が低いものから高いものへ、さらに平均他者評価が低いものから高いものへと並べて決定した。



図5 自己評価と平均他者評価の例（本人の許可を得て掲載）

(注) 左（図4左のID8）の自己評価は5、平均他者評価は6.8で、自己評価は信頼区間の下限の下だった。右（図4左のID13）の自己評価は6、平均他者評価は6.1で、信頼区間に入っていた。

表2 平均他者評価、他者評価の中央値を従属変数とした回帰分析結果

	従属変数		
	平均他者評価		他者評価の中央値
	モデル 1	2	
男性ダミー	0.019	- 0.001	0.006
自己評価	0.517 *	0.720 *	0.728 *
最低推測		- 0.106	- 0.074
最高推測		- 0.209	- 0.273
決定係数	0.271	0.322	0.319

(注) $N = 20$ 。値は標準化係数。* $p < .05$ 。

$1.84 + 0.65 \times$ 自己評価となった。平均他者評価と自己評価の相関係数は、0.520 だった ($N = 20$ 、 $p = 0.019$)。

3.2 信頼区間

では、自己評価はどれくらい他者評価の平均の信頼区間に入ったのだろうか。20名の自己評価のそれぞれについて、他者評価の平均の95%信頼区間を求め、自己評価が低いものから並べたのが図4左である。

信頼区間に入ったのは、20名中8名だった。範囲外だった12名のうち、6名は自己評価が信頼区間の上限より上で、6名は下限より下だった。したがって、自己評価と平均他者評価は完全には一致しないが、ゆるやかに関連していることが示唆された。

図5は、自己評価と平均他者評価の2つの例である（本人の許可を得て掲載）。図5左は自己評価が信頼区間の下限の下であったので、本人が過少評価していた。図5右は、自己評価が信頼区間に入っていたので、容姿を正確に把握していた。

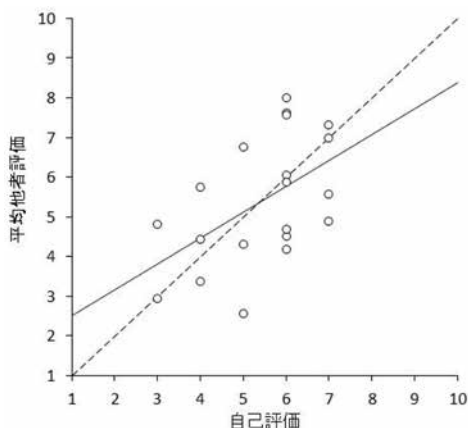


図6 自己評価と平均他者評価の回帰直線

(注) $N = 20$ 。実線は単回帰直線で、平均他者評価 = $1.84 + 0.65 \times$ 自己評価。破線は45度線で、回帰直線と自己評価5.33で交わる。

3.3 平均のグループ別比較

グループ別に、平均他者評価を比較すると違いがあるだろうか。被評価者の性別で比較すると、男性の平均他者評価の平均が5.69、女性だと5.32だった（分散分析で差なし、 $N = 20$ 、 $p = 0.667$ ）。

自己評価グループ別の比較は、図4右となった。自己評価が上がるほど、平均他者評価の平均も上昇した。ただし、分散分析では差がなかった（ $N = 20$ 、 $p = 0.205$ ）。

3.4 回帰分析

他の変数の効果を統制するため、平均他者評価を従属変数と、自己評価を独立変数として、回帰分析を実施した。表2がその結果である。モデル1では独立変数である自己評価の効果と、統制変数として性別の効果が検討されている。また、もしかしたら自己評価よりも、最低推測や最高推測のほうが平均他者評価を予測できるかもしれない。そこで、モデル2では、さらに最低推測と最高推測で統制する。

どちらのモデルでも、自己評価のみが有意な（正の）効果を持った（モデル1で係数0.517、モデル2で0.720）。つまり、自己評価が高い人ほど、他者からの平均他者評価が有意に高かった。性別、最低推測、最高推測は有意な効果を持たなかった。

なお、（表では省略したが）モデル1と2とで、決定係数に有意な差はなかった（決定係数はそれぞれ0.271、0.322）。そのため、モデル1で十分であった。したがって、平均他者評価を予測するには自己評価のみを用いればよく、さきほどの「平均他者評価 = $1.84 + 0.65 \times$ 自己評価」という単回帰式で求められることが分かった。この回帰式によれば、自己評価が5.33のとき45度線と交わる（図6）。そのため、自己評価が1下から10上のうち、1から5まででは自己評価のほうが平均他者評価より低く、過少評価される。6以上では、かえって自己評価のほうが高く、過大評価

となっていた。

自己評価、最低推測、最高推測は、分布で確認したとおり関連しあっていた。回帰分析では、VIF がもっとも高いもので、3.9 だった。

結果の頑健性をチェックするために、他者評価の平均ではなく中央値を従属変数として分析した(表2)。それでも、同じ結果となった。

4 考察

4.1 分析結果の要約

- (1) 分布より、容姿の自己評価と平均他者評価に平均の違いはなかった。
- (2) 他者評価の平均の信頼区間より、被評価者のうち半分ほど(20名中8名)が、95%信頼区間に入っていた。したがって、自己評価と平均他者評価は完全には一致しないが、ゆるやかに関連していることが示唆された。
- (3) 平均他者評価を従属変数とした回帰分析より、性別、最低推測、最高推測は効果を持たず、自己評価だけが有意な(正の)効果を持った。その結果、平均他者評価を予測するには自己評価のみを用いればよく、自己評価が高い人ほど平均他者評価が高かった。
- (4) ただし、回帰式より、自己評価が低い人ほど自己評価が平均他者評価より過少評価に、高い人ほど過大評価になっていた。

以上から、自己評価と平均他者評価は完全に一致するわけではないが、おおむね自己評価で平均他者評価を予測できることが分かった。つまり、人びとはおおむね正確に自分の容姿を把握していた。その結果、「容姿の自己評価が高い人ほど、他者評価が高いだろう」という仮説の検証結果は、以下となる。

仮説の検証結果。容姿の自己評価が高い人ほど、平均他者評価が有意に高かったので、仮説は支持された。ただし、自己評価が低い人ほど自己評価が平均他者評価より過少評価に、高い人ほど過大評価になっていた。

4.2 リサーチ・クエスチョンへの回答

ここから、「容姿の自己評価は他者評価と一致するのか」というリサーチ・クエスチョンにたいして、つぎのように回答することができる。容姿は主観的評価でも十分に測定できたので、自計式調査での測定である程度信頼できることが分かった。

リサーチ・クエスチョンへの回答。容姿の自己評価は、平均他者評価とおおむね一致する。したがって、容姿の測定には、自計式調査でもある程度信頼できるので、自計式調査で容姿測定のデータ収集をできる可能性が示された。

たとえるなら、容姿とはインディアン・ポーカーに似ているかもしれない。インディアン・ポーカーでは、他人のカードの強さ（数字）はすべて分かるのに、自分のカードの強さだけ分からない。そうした中で、他人との会話やコミュニケーションをとおして自分のカードの相対的な強さを推測して、勝負するかどうかを決定する。容姿についても同様に、周囲の人からの扱いをもとに「自分の容姿の水準はこれくらいかな」と推しはかる必要がある。

とはいえ、この論文の分析結果によれば、そうして得られた自己評価は、他者からの評価と完全には一致しないものの、おおむね近似していることが分かった。人びとは、容姿というインディアン・ポーカーを日々プレーし、自分への評価を修正しつづけているのかもしれない。

4.3 今後の課題

- (1) 今後は、自計式調査で容姿のデータを幅広く蓄積することで、容姿の規定構造や、容姿の違いの帰結を解明することが期待される。
- (2) そうしてデータ収集したうえで、容姿の違いをどのように理論的に説明できるかが、そのつぎの課題となろう。谷本（2015）は、美容医療や美容整形を行なうことを、「外見資本」への投資と捉える。Hakim（2011）は *erotic capital* への、小林（2017）は美容資本への投資として概念化している。
- (3) この論文の分析結果を調査方法論の視点から考えるなら、自計式調査の信頼性が示唆され、可能性がより開かれたといえる。そこで、今後は（容姿に限らず）行動や心理などについて自計式調査で主観的評価を行なったとき、それらが他者からみた客観的評価とどれくらい一致するのかしないのか、それはなぜかを検討することが必要だろう。

【謝辞】

本研究は、JSPS 科研費 JP15H01969（2015～18年度、小林盾代表）の助成を受けたものです。自己評価のデータ収集について、成蹊大学研究倫理委員会から承認を得ました（評価についての研究、SREC15-05）。大崎裕子氏がデータ整理で協力してくれました。自己評価、他者評価ともに、データ収集への協力者のおかげで分析ができました。片瀬一男氏、川端健嗣氏、佐藤嘉倫氏、七條達弘氏、永吉希久子氏、浜田宏氏、山田昌弘氏から有益なコメントをいただきました。記して感謝します。

【文献】

- Bjerk, D., 2009, "Beauty vs. Earnings: Gender Differences in Earnings and Priorities over Spousal Characteristics in a Matching Model," *Journal of Economic Behavior & Organization* 69: 248-59.
- Buss, D. M. 1994, *The Evolution Of Desire: Strategies of Human Mating*, Basic Books.
- Campbell, A., P. E. Converse, and W. L. Rodgers, 1971, *Quality of American Life Questionnaire*, Inter-university Consortium for Political and Social Research 3508.
- Etcoff, N., 1999, *Survival of the Prettiest: The Science of Beauty*, Anchor. (= 2000、木村博江訳、『なぜ美人ばかりが得をするのか』草思社。)

- Hakim, C., 2011, *Erotic Capital: The Power of Attraction in the Boardroom and the Bedroom*, Basic Books.
- Hamermesh, D. S., 2006, "Changing Looks and Changing "Discrimination": The Beauty of Economists," *Economics Letters* 93: 405-412.
- Hamermesh, D. S., 2011, *Beauty Pays: Why Attractive People Are More Successful*, Princeton University Press. (= 2015、望月衛訳『美貌格差：生まれつき不平等の経済学』東洋経済新報社。)
- Hamermesh, D. S. and J. E. Biddle, 1994, "Beauty and the Labor Market," *American Economic Review* 84(5): 1174-94.
- Hatfield, E. and S. Sprecher, 1986, *Mirror, Mirror...: The Importance of Looks in Everyday Life*, State University of New York Press.
- 小林盾、2017、『ライフスタイルの社会学：データからみる日本社会の多様な格差』東京大学出版会。
- 小林盾・谷本奈穂、2016、「容姿と社会的不平等：キャリア形成、家族形成、心理にどう影響するのか」『成蹊大学文学部紀要』51: 99-113。
- 谷本奈穂、2008、『美容整形と化粧の社会学：プラスチックな身体』新曜社。
- 谷本奈穂、2015、「美容：美容整形・美容医療に格差はあるのか」小林盾・山田昌弘編『ライフスタイルとライフコース：データで読む現代社会』新曜社。
- Umberson, D. and M. Hughes, 1987, "The Impact of Physical Attractiveness on Achievement and Psychological Well-Being," *Social Psychology Quarterly* 50(3): 227-36.
- 山田昌弘・白河桃子、2008、『「婚活」時代』ディスカヴァー・トゥエンティワン。
- Wolf, N., 1991, *The Beauty Myth: How Images of Beauty Are Used Against Women*, Anchor.