

Xavier Caroff, Maud Besançon,

Variability of creativity judgments.

Learning and Individual Differences,
Volume 18, Issue 4, 2008, Pages 367-371.

1. Introduction

- ✚ 近年、創造性の定義を統一する傾向にある
 - 創造性とは、新しく、かつ、文脈に合った生成物を実現する能力のこと (e.g., Amabile, 1996, Lubart, 1994, Sternberg and Lubart, 1995)
- ✚ 様々なドメインの生成物の創造性評価には、CAT (Consensual Assessment Technique; Amabile, 1982, Amabile, 1996, Hennessey and Amabile, 1999) がよく使われてきた
 - CAT では、その生成物のドメインに精通する専門家が1つずつ創造性を評価する
 - ✧ 特に創造性の評価基準を与えられることはなく、主観的な理解のもとで評価させている
- ✓ 生成物、評価課題、評価者自身のどういった特徴が創造性評価に影響しているのかを検討する研究が必要 (Amabile, 1996)

<本研究の主な目的>

- ✓ 広告の創造性評価の個人差は、以下の差異にどの程度起因するのかを検討すること
- ① 評価者が創造性の2つの基本的な要素（独創性と適切性）に依存する程度と、これらをどのように統合して創造性の全体的な判断を行うかという個人差
- ② 評価者の特性（能力としての創造性の個人差など）
- ③ 創造性について評価者が受ける教示の違い

1.1. Information integration in creativity judgments

- ✚ 筆者の知る限り Runco and Charles (1993)が、独創性と適切性がどのように創造性評価に貢献するかを検討した唯一の研究
 - 拡散的思考テストで生成された生成物から3種類のアイデアセットを作成した
 - ✧ その結果、創造性評価は、独創的で適切なアイデアの割合が増えた場合と、独創的だが不適切なアイデアの割合が増えた場合の両方で、向上することが示された
 - ✧ 逆に、適切だがありふれたアイデアの割合が増えると、創造性評価は低下する
- ✚ 広告のような現実の生成物の創造性評価で、基本的な2つの要素（独創性と適切性）を扱った研究は非常に少ない
 - 広告の創造性を評価する際に両方の要素を変化させた場合、評価者によって両方の特性に依存する程度が異なるのではないか、という疑問が残っている

- ✓ 本研究は、Runco and Charles (1993)の提言に基づいて、評価者が独創性と適切性に関する情報を統合して創造性を評価するメカニズムを調査する
 - 参加者は、専門家によって事前に独創性と適切性が評価され、2つの要素が系統的に変化する広告セットに対して創造性を評価した

<仮説>

- ✓ 独創性のレベルが高くなるにつれて創造性評価も高くなる
- ※ 適切性も広告の創造性評価を調整すると考えられるが、仮説を立てるのは難しい
 - 定義上、適切性も兼ね備えるべきだが、陳腐になるほどの過度の適切性ではない
 - 独創性と適切性の情報を統合して判断されるなら、この2つの要素の交互作用が観察されるはずである

1.2. Individual differences in judgments related to the creative ability of the judges

- ✚ Hood (1973)は、Unusual Uses Testで測定された評価者の独創性と、同じテストの別の形式で得られた他者の生成物に対する独創性評価との関連を調べた
 - その結果、独創性の低い評価者の創造性評価は生成物の独創性レベルに大きく影響されるのに対して、中程度から高い独創性を持つ評価者は生成物の創造性をあまり認識せず評価平均が低くなることが明らかになった
 - ① これらの結果から、独創性の高い評定者は、創造性について非常に限定的な概念を持ち、非常に独創的な生成物のみが創造的であると考えている可能性が示唆された
- ✓ 本研究は、創造性評価に、評価者の独創性と生成物の独創性による交互作用があることを示す先行研究の結果を再現する
- ※ しかし、評価者の独創性が広告の適切性と交互作用を示すことを期待する理論的背景はない

1.3. Individual differences in judgments related to instruction

- ✚ CATは、評価者が評価する生成物の分野の専門知識を持っていれば、生成物の創造性を自発的に認識すると仮定している (Amabile, 1982, Amabile, 1996)
 - 様々な実験条件下で、通常、信頼性が高いことが示されている (e.g., Amabile, 1996, Baer, 1994, Baer et al., 2004, Hennessey, 1994, Hennessey and Amabile, 1999)
 - しかし、評価者の専門性の高さが創造性評価の個人差となっている
 - ① 専門家の評価者間の合意は、非専門家よりも劣っていることがある (Dollinger et al., 2004, Hickey, 2001)
 - ✧ 専門家は、その生成物のドメインを通して培った経験をもとに、主観的な創造性の理解を発展させてきたと解釈できる
 - ② (驚くことではないが)非専門家の評価者間合意は少々低い
 - ✧ 評価者間合意を高めるために、Dollinger and Shaffran (2005)は非専門家に創造性を評価してもらう前に、評価者の評価を校正するトレーニングを行うことを提案した
 - ✧ トレーニングはCATで推奨されていないが、創造性評価時の教示やトレーニングの効果に関して一般的な疑問を生じさせる

<仮説>

- ✓ 創造性評価の基準となる要素が明示されていれば、評価者が独創性と適切性の両方を考慮する傾向にある
 - 創造性の概念的枠組みが与えられていない状況と比較して、創造性の規範的な定義と明確な評価基準が提示され、さらに創造性を評価する前に両方の基準を評価するように訓練された場合、創造性の評価を決定する上で独創性と適切性のそれぞれの重要性が高まるはず
 - この仮説が正しければ、創造性評価において、3つの評価状況は、広告の独創性と適切性の両方の特性と交互作用する

2. Method

- ✚ 3つの参加者間要因（性別、評価者の独創性レベル高低、3つの実験状況）と2つの参加者内要因（広告の独創性レベル低中高、広告の適切性レベル低中高）の混合計画
 - 性別は、それが創造性評価に影響しているかを確認するために含めた

2.1. Subjects

- ✚ パリ大学の学生 95 名（平均年齢 27.7 歳， $SD = 4.9$ ），1/3 は創造性の講義を受講している

2.2. Material

2.2.1. Test of divergent thinking

- ✚ 参加者の創造性の能力を測るために、拡散的思考テストの Torrance Tests of Creative Thinking (TTCT; Torrance, 1976) に回答してもらった
 - 段ボール箱の変った使い方を 10 分以内にできるだけ多く考え出す課題

2.2.2. Construction of a set of advertisements to be evaluated

- ✚ 自動車の広告を扱った
 - 一般の新聞に掲載されていて参加者にとって身近なため
- ✚ 190 の広告を集め、各広告に対して広告代理店に勤める専門家（女性 15 名，男性 9 名；25 歳～50 歳）が独創性，適切性，創造性を 0（全く〇〇がない）～7（非常に〇〇が高い）で回答した
 - 独創性，適切性，創造性の専門家の平均値を算出
- ✚ 専門家の評価をもとに、実験刺激として、15 の広告からなる評価セットを作成
 - 評価の統合について調査するために、Anderson (1982a,b,1996) が提唱する方法ガイドラインに従い、評価平均の分布に基づいて、2 種類の広告を用意した
 - ① 独創性と適切性が両方存在し、それぞれが低・中・高のレベルで変化している広告 9 つ
 - ② 不適切だが独創性が低・中・高のレベルで変化している広告 3 つ，普遍的だが適切性が低・中・高のレベルで変化している広告 3 つ
 - 15 の広告は、評価者間の一致が認められた（独創性 $\alpha = .92$ ，適切性 $\alpha = .76$ ，創造性 $\alpha = .94$ ）

2.3. Procedure

1. 拡散的思考タスク

- 参加者の独創性を測るために実施
- タスクで生成したアイデアの生成頻度をもとに、中央値より低い参加者は独創性低群、中央値より高い参加者は独創性高群に群分けした

2. 実験タスク

- 状況③は創造性講義の受講者が、状況①②はランダムに割り当てられた
 状況① 定義を示さずに創造性を評価する（自身の潜在的概念理解に頼る）
 状況② 定義を示した上で創造性を評価する
 状況③ 創造性に関連する要素を評価するトレーニングを実施した後に創造性を評価する
 - ◇ 各状況には男女各 12 名、独創性高群と低群が半数ずつになるよう割り当てられ、独創性スコアが中央値に近い参加者は除いた
- 広告は固定のランダム順で 4 回通り提示され、毎回創造性を評価してもらった
 - ◇ 1(創造的でない)~7(非常に創造的)で評価
 - ◇ 独立変数は、各広告に対する 4 回分の創造性評価の平均

3. Results

- 📊 3つのレベルの独創性（低、中、高）と3つのレベルの適切性の各組み合わせに対応する9つの広告のみを分析対象とした
 - 参加者の創造性評価は、3つの参加者間要因（性別、2つの評価者の独創性、3つの実験状況）と2つの参加者内要因（3段階の広告の独創性と3段階の広告の適切性）を用いた混合計画のANOVAで分析した

- 📊 結果は Table1 に示した通り

<主効果>

- 男女に差は見られず ($F(1,60)=.05$ NS), 他の要因との交互作用もなかった
- 独創性レベルが異なる2群の評価者間に有意な創造性評価の差はなかった ($F(1,60)=.04$ NS)
- 唯一の有意な主効果は広告の特性に関するものだった
 - ◇ 創造性評価は、広告の独創性が高いほど ($F(2,120)=243.61, p < .001, \text{partial } \eta^2=.802$), 適切性が高いほど ($F(2,120)=3.80, p < .05, \text{partial } \eta^2=.060$) 高かった

<交互作用>

- 広告の独創性と適切性の交互作用が有意だった ($F(4,240)=9.97, p < .001, \text{partial } \eta^2=.143$)
 - ◇ Figure 1 は参加者が、創造性に関連する2つの要素を特定の方法で統合していることを示唆している
 - 創造性を評価する際、広告の適切性が低い場合は、適切性が中程度や高い場合よりも、広告の独創性に影響を受けていることを示している

※ しかし、この交互作用グラフは、Anderson (1982a,b,1996)が提唱した代数的統合モデルとは一致せず、平行な曲線で構成されているわけでもなく、特徴的な線形ファン効果を示しているわけでもなかった

➤ このため、不適切で（または普遍的で）独創性（適切性）レベルが変化する6つの広告について得られた創造性評価を用いて、統合のメカニズムを深堀することはできなかった

Table 1
Results of the mixed factorial ANOVA

Sources	Effect		Error		F	Sig.	Partial η^2
	df	MS	df	MS			
Gender (G)	1	.232	60	4.406	.053	.819	.001
Raters Originality (RO)	1	.162	60	4.406	.037	.849	.001
Situations (S)	2	1.797	60	4.406	.408	.667	.013
Advert. Originality (AO)	2	511.337	120	2.099	243.610	<.001	.802
Advert. Appropriateness (AA)	2	3.781	120	.996	3.796	.025	.060
G×RO	1	16.772	60	4.406	3.806	.056	.060
G×S	2	5.742	60	4.406	1.303	.279	.042
RO×S	2	6.208	60	4.406	1.409	.252	.045
G×AO	2	1.370	120	2.099	.653	.523	.011
RO×AO	2	7.091	120	2.099	3.378	.037	.053
S×AO	4	1.569	120	2.099	.747	.562	.024
G×AA	2	.132	120	.996	.133	.876	.002
RO×AA	2	.193	120	.996	.194	.824	.003
S×AA	4	2.641	120	.996	2.652	.036	.081
AO×AA	4	11.420	240	1.145	9.971	<.001	.143
G×RO×S	2	3.356	60	4.406	.762	.471	.025
G×RO×AO	2	.151	120	2.099	.072	.930	.001
G×S×AO	4	.533	120	2.099	.254	.907	.008
RO×S×AO	4	2.219	120	2.099	1.057	.381	.034
G×RO×AA	2	.995	120	.996	.999	.371	.016
G×S×AA	4	.408	120	.996	.409	.802	.013
RO×S×AA	4	1.795	120	.996	1.803	.133	.057
G×AO×AA	4	1.046	240	1.145	.913	.457	.015
RO×AO×AA	4	.564	240	1.145	.493	.741	.008
S×AO×AA	8	1.159	240	1.145	1.012	.427	.033
G×RO×S×AO	4	3.715	120	2.099	1.770	.139	.056
G×RO×S×AA	4	.967	120	.996	.971	.426	.031
G×RO×AO×AA	4	1.455	240	1.145	1.271	.282	.021
G×S×AO×AA	8	1.012	240	1.145	.883	.531	.029
RO×S×AO×AA	8	1.170	240	1.145	1.022	.420	.033
G×RO×S×AO×AA	8	.779	240	1.145	.680	.709	.022

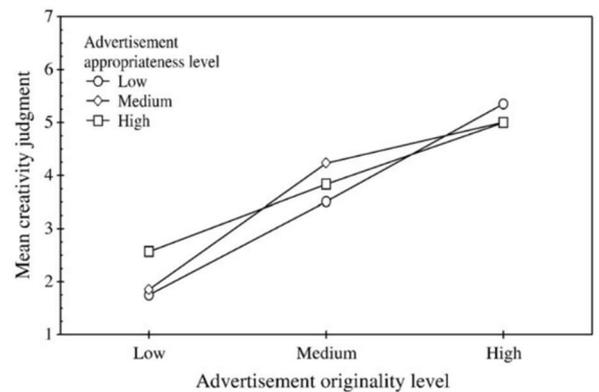


Fig. 1. Mean creativity judgments for advertisements of different levels of originality and different levels of appropriateness.

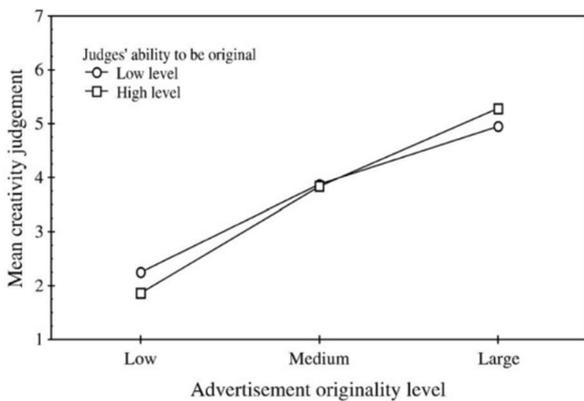


Fig. 2. Mean creativity judgments for advertisements of different levels of originality and judges ability to be original.

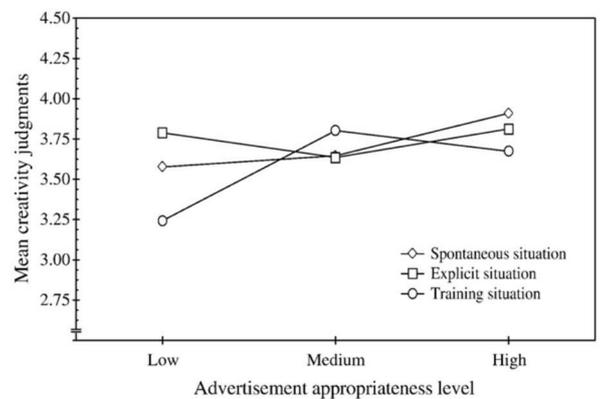


Fig. 3. Mean creativity judgments for advertisements of different levels of appropriateness and for different judgment situations.

- Hood (1973)が示唆しているように、評価者の独創性と広告の独創性の交互作用が有意だった ($F(2,120)=3.38, p < .05, \text{partial } \eta^2=.053$)
 - ◇ しかし予測と異なり、創造性評価に広告の独創性が与える影響は、独創性の高い評価者の方が低い評価者よりも大きかった (Figure 2)
- 参加者の独創性と広告の適切性の交互作用は有意ではなかった ($F(2,120)=.19 \text{ NS}$)
- 参加者の独創性を二分した両群においても、広告の適切性が低い場合は、中程度や高い場合に比べて、広告の独創性に影響されることを確認した
 - ◇ 独創性低群は ($F(4,120)=5.54, p < .001, \text{partial } \eta^2=.156$)
 - ◇ 独創性高群は ($F(4,120)=4.89, p < .01, \text{partial } \eta^2=.140$)
- 予測と異なり、創造性評価に広告の独創性が与える影響は、3つの評価状況で変化しなかった ($F(4,120)=0.75 \text{ NS}$)
- 一方、広告の適切性と評価状況の間には有意な交互作用が見られた ($F(4,120)=2.65, p < .05, \text{partial } \eta^2=.081$)
 - ◇ この交互作用 (Figure 3) を、トレンド分析でさらに検討した
 - 定義を示さずに創造性を評価してもらった場合 (状況①)、広告の適切性が高いほど創造性評価が高いことが示された ($F(1,60)=2.65, p < .05, \text{partial } \eta^2=.042$)
 - 曲線的関係は示さなかった ($F(1,60)=.39 \text{ NS}$)
 - 定義を与えて創造性を評価してもらった場合 (状況②)、線形・曲線的関係の両方が有意ではなかった ($F(1,60)=.02 \text{ NS}; F(1,60)=1.13, p > .05$)
 - 創造性に関連する要素で評価するトレーニングを行った場合 (状況③)、線形・曲線的関係の両方が有意だった ($F(1,60)=7.92, p < .01, \text{partial } \eta^2=.117; F(1,60)=4.93, p < .05, \text{partial } \eta^2=.076$)

4. Discussion and conclusion

- 📌 本研究の目的① 独創性と適切性が、創造性とどのように結びついているのか理解すること
 - 先行研究 (Hood, 1973; Runco & Charles, 1993) から、広告の独創性が高いほど、創造性評価が高くなることは驚くべき結果ではない
 - しかし、適切性の効果に関する結果はやや複雑であるが、生成物の適切性が独創性のレベルに基づく創造性の評価を調整する可能性があることを発見した
 - ◇ 広告の独創性が創造性評価に与える影響は、適切性が低い場合の方が、適切性が中程度や高い場合よりも大きい (適切性は、独創性が創造性評価に与える影響を減少させる)
 - この交互作用は、独創性スコアで二分した両群でも観察された
 - 評価デザイン (状況) を変化させても、広告の独創性が創造性評価に与える影響は変わらない
 - ◇ 評価者の自発的な創造性評価に独創性が基づいている可能性があり、この基準はどんな教示やトレーニングによっても修正されないと考えられる
 - しかし、広告の適切性が創造性評価に与える影響は評価状況で変化する

- ◇ 創造性の概念的枠組みを与えずに評価すると、創造性は適切性と直線的な関係にあった
- ◇ 創造性の概念的枠組みが与えられて評価すると、適切性は創造性と無関係だった
- ◇ 創造性評価を 2 つの要素で評価する訓練を受けると、適切性が高すぎず低すぎない広告が最も創造的であると評価する
 - この結果は、適切性の閾値が創造性評価に関係している可能性を示唆し、この閾値を超えると、独創性とは無関係に創造性がないと評価されることになる
 - この解釈は、今後の研究で検証されるべき（今回の結果がドメインを超えて一般化されるのか、ドメイン固有のものなのかも検討されるべき：芸術分野では、独創性がより重視され、他の領域では独創性よりも適切性の方が影響力を持つ可能性がある）

✚ 本研究の目的② 評価者の能力としての創造性が創造性評価に影響を与えるか検証すること

- 広告の独創性が創造性評価に及ぼす影響は、独創性が高い評価者の方が、低い評価者よりも大きいことが明らかになった
- ◇ Hood(1973)の結果に基づく、独創性が高い評価者は創造性を非常に限定的に捉えており、かなり独創的な生成物のみが創造的であると考えているのではないかという仮説を裏付けるものではない
- ◇ しかし、独創性が高い評価者は、より独創性に依拠して創造性を差別化していることが確認された
 - 評価者の適切な生成物を生み出す能力を評価し、創造性評価において、その能力と適切性や独創性との交互作用を調べることも興味深い

✚ 結論として、本研究は、創造性評価における情報統合のメカニズム、特に適切性が創造性評価に及ぼす影響を研究することの重要性を強調する

- また、創造性評価にはばらつきがあることが示され、評価者間の個人差をより体系的に研究することを正当化できた(Hennessey, 1994)