

表情解読の正確さを測定する日本人用課題遂行型テストの作成

——表情の強度と表情解読の正確さとの関連に着目して——

22001FPM 志水 勇之進

キーワード：表情解読・表情の強度・課題遂行型テスト

問題と目的

対人コミュニケーションにおいて、受け手の解読の正確さを高めることは円滑な関係をもたらす要因の一つである(大坊, 2006)。表情は送り手の感情を表すという意味において、受け手にとって有用な情報源となる(Ekman, 2003 など)。表情を介して伝達されるメッセージを正確に解読することは、人間関係を良好に保つために重要であるとの指摘もある(Hall, 1980 など)。解読の正確さをパフォーマンスに基づき測定する課題遂行型テストは国外には存在するが、いずれも刺激人物が欧米人であるため、コミュニケーション行動における文化的差異から(Elfenbein & Ambady, 2002 など)、日本人の解読の正確さを測定するテストとして最適であるとはいえない。さらに、これまで表情の強度が最大である刺激が用いられてきたが、日本人は感情を抑制したうえで表出する傾向がある(Ekman & Friesen, 1975)。以上、本論文の第一の目的は、日本人を刺激人物とし、また表情刺激の強度を操作したうえで、表情解読の正確さを測定する日本人用課題遂行型テストを作成することである。第二の目的は、強度が異なる表情における解読の特徴を明らかにするため、基本感情に対応する各表情において、表情の強度と表情解読の正確さとの関連について検討を行うことである。

研究1 目的

日本人を刺激人物とし、表情刺激の強度を操作したうえで、表情解読の正確さを測定する日本人用課題遂行型テストの作成を試みる。

実験1 方法

実験参加者 大学生40名, 社会人31名

表情刺激 AIST 顔表情データベース 2017 (Fujimura & Umemura, 2017) のうち男性モデル4名, 女性モデル4名の正面顔を用いた。各モデ

ルの怒り(開口・閉口)、嫌悪(開口・閉口)、恐れ、喜び、悲しみ、驚きの8表情、計64刺激のうち、Face Reader や FACS コーディングによって当該表情として妥当であると判断された37刺激を選定した。次に、抽出された刺激を強度100、中立表情を強度0とし、強度10%ごとにモーフィングした370刺激と、各刺激人物の中立表情7刺激を合わせた、377刺激を使用した。

実験手続き 実験室のモニターの画面中央に表情刺激を呈示し、7つの選択肢(怒り、嫌悪、恐れ、喜び、悲しみ、驚き、無表情)の中から、呈示された表情刺激に最も当てはまると思われるものを1つ選択させた。

実験1 結果

正答率が10%以下、90%以上であった刺激を除外した。次に、(a)強度と正答率の関係が線形にならない刺激は除外する、(b)表情カテゴリー間の刺激数や、各表情カテゴリーにおける強度、刺激人物の数のバランスを考慮して選定する、という基準に則り、刺激を選定した。その結果、合計114刺激(中立7、怒り17、嫌悪19、恐れ21、喜び14、悲しみ21、驚き15)となった。

実験2 方法

実験参加者 20代から60代の男女各100名以上を対象に調査会社を通じてオンライン実験を行い、1140名から回答を得た。Directed Questions Scale への誤答者や回答時間が極端に長いまたは短かった回答者を除外し、554名を分析の対象とした。

表情刺激 実験1の結果選定された114刺激。

実験手続き 実験1と同様の手続きでオンライン実験を行った。

実験2 結果

正答率が10%以下の刺激、90%以上の刺激、合計得点との点双列相関係数が.20以下であった刺

激を除外したのち、テストの1次元性の検討のため、カテゴリカル因子分析を行った。なお、ここまでの手続きで中立表情の6刺激は除外されたため、残った1刺激についても除外した。固有値の減衰状況や因子負荷量などから1次元性が認められなかったため、GERT作成時の手続きになら（Schlegel et al., 2014）、表情カテゴリーごとにカテゴリカル因子分析を行った。嫌悪表情において因子負荷量が低かった刺激を除外した結果、すべての表情カテゴリーで1次元性が認められた。次に、各表情カテゴリーにおいて項目反応理論を適用した分析を行い、項目パラメータの推定を行った。その結果、識別力が極端に低い刺激はみられなかったが、表情カテゴリーごとに(a)困難度の値が同等の刺激が複数ある場合、除外対象とする、(b)その場合、原則として識別力が低い刺激を除外対象とするが、表情の強度、刺激人物のバランスを考慮して選定する、という手順で選定を行った。以上の結果採択された78刺激（怒り14、嫌悪14、恐れ14、喜び14、悲しみ14、驚き8）について、 ω 係数を算出したところ.85であったため、日本人の表情解読の正確さを測定することが可能であると判断した。このテストは、Facial Expressions of Emotion Test focused on Intensity — using the AIST facial expression database 2017 — (FEET-I) と命名した。

実験3 方法

実験参加者 大学生および大学院生 85名

テスト刺激 ①FEET-I：実験2の結果採択された78刺激。②GERT-S：Schlegel & Scherer (2015)による感情解読テスト、計42刺激。③アジア版RMET：Adams et al. (2010)で用いられた、精神状態の解読テスト、計36刺激¹。

実験手続き 実験1と同様の手続きで、強制選択法による解読実験を行った。

実験3 結果

FEET-Iの収束的妥当性を検討するために、GERT-S、アジア版RMETとの関連について検

討した結果、GERT-Sとの間では有意な正の相関を示したが ($r = .472, p = <.001$)、アジア版RMETとの間には有意な相関はみられなかった ($r = .163, ns$)。

研究2 目的

各表情について強度と解読の正確さに線形関係がみられるかどうかを検討する。したがって、強度のバリエーションが豊富な刺激を用いた、研究1の実験1で得たデータについて分析を行う。

研究2 結果

中心化した各刺激の強度、強度の2乗項、強度の3乗項を独立変数、各刺激に対する平均正答率を従属変数として、表情カテゴリーごとに階層的重回帰分析を行った。その結果、怒り、嫌悪、恐れ、悲しみ ($\text{adj } R^2 = .492 \sim .897, p < .001$) については、2次モデルが採用され、強度の2乗項の効果がいずれにおいても有意であった ($B = -.010 \sim -.016, p < .001$)。つまり、表情の強度が強くなるにつれ解読の正確さが上昇するが、ある時点で正確さが低下することを示している。喜びと驚きについては、3次モデルが採用され（それぞれ $\text{adj } R^2 = .742, p = .019$; $\text{adj } R^2 = .772, p < .001$ ）、強度の3乗項の効果がそれぞれ有意であった（それぞれ $B = .00027, p = .019$; $B = .00037, p < .001$ ）。つまり、強度が強くなるにつれ、解読の正確さが上昇するが、ある時点で正確さが低下し、そしてある時点で正確さが再び上昇または停滞することを表している。

総合考察

FEET-Iの作成に関して、テストの内的一貫性が認められたうえ、感情解読の正確さを測定するGERT-Sとは有意な正の相関を示したことから、FEET-Iは、感情解読の正確さを測定するテストとして妥当であるといえる。また、表情の強度と解読の正確さとの関係を検討した結果、すべての表情について線形項は最終モデルとして採用されなかった。つまり、表情の強度が強いほど正確に解読されるわけではないといえる。日本の社会的文化的文脈における生態学的妥当性を確保するためにも、強度が異なる表情における解読の特徴について、さらに検討していく必要があるだろう。

¹ 吉川左紀子（京都大学）・野村光江（関西看護医療大学）によって作成されたものを使用した。