

表情認知課題と視点取り課題を用いた自閉傾向の予測

岩内 厚大[†] 田中 宏季[†] 中村 哲[†]

[†] 奈良先端科学技術大学院大学 情報科学領域

E-mail: †{iwauchi.kota.ig9,hiroki-tan}@is.naist.jp

あらまし 自閉スペクトラム症は、精神科医が診断基準に基づき、診断を行なっている。臨床診断の精度を向上させるために、自閉特有の行動指標は必要とされている。我々はこれまで、表情認知時の眼球運動を用いて自閉スペクトラム症の鑑別を試みる研究を行ってきた。自閉傾向の高い方は心の理論の程度が低いことが知られており、またこれによって社会的視点取得を苦手としている。本研究では、そこに焦点を当てた課題として、空間的視点取得課題の眼球運動を用い、表情認知時の眼球運動と合わせることで、さらに高精度で自閉傾向を予測できることを提案する。22-35歳の27名のデータを収集し、自閉傾向の指標である対人応答性尺度(SRS-2)を予測した。結果として、二つの課題から得た眼球運動を用いることでRMSEが24.40で予測を行うことができ、表情認知時のみの眼球運動を用いるときに比べ、RMSEで0.92上回る性能で予測可能であることを確認し、予測値と真値との無相関検定において $p < 0.05$ で有意であることを確認した。

キーワード 自閉スペクトラム, 眼球運動, 機械学習

Predicting Autism Using Facial Emotion Identification Test and Visual Perspective Taking

Kota IWAUCHI[†], Hiroki TANAKA[†], and Satoshi NAKAMURA[†]

[†] Division of Information Science, Nara Institute of Science and Technology

E-mail: †{iwauchi.kota.ig9,hiroki-tan}@is.naist.jp

Key words Autism, Eye-movements, Machine Learning

1. ま え が き

自閉スペクトラム症は、精神科医が診断基準に基づき、診断を行なっている [1]。臨床診断の精度を向上させるために、自閉特有の行動指標は必要とされている。自閉群は表情認知の際に、健常群と違う視線走査を行うことが知られており、我々はこれまで、表情認知時の眼球運動を用いて自閉スペクトラム症の評価を試みる研究を行ってきた [2]。我々はこれまでそこに焦点をあてて調査を行ってきた。また、表情認知課題時の眼球運動を用いて自閉傾向を予測できることが知られている [3]。しかしながら、自閉傾向と関係がある能力はこれだけではない。そのうちのひとつとして我々は心の理論に焦点をあてる。

心の理論とは、他者の感情や思考、意図などの心的状態に対する推定、および他者は自己と異なる心的状態が存在していることを理解することを指す。自閉スペクトラム症の方は他者の心的状態を予測することが苦手なことが知られている [4]。心の理論と社会的な視点認識は関連があり、また、社会的視点

認識と空間的な視点認識の関連も知られている [5]。

以上の点を考慮し、我々は表情認知課題時の眼球運動と視点取り課題時の眼球運動を用いて、複合的に認知活動を測定することにより、より高精度に自閉傾向を予測できる手法を提案する。具体的には自閉傾向の簡易指標である対人性応答尺度(SRS-2)を予測対象とし、表情認知課題時の眼球運動による予測、視点取り課題時の眼球運動による予測、表情認知課題、視点取り課題時の眼球運動を複合的に用いたときの予測結果を比較した。結果として、複合的に眼球運動を用いたときの予測結果が決定係数で0.121、スピアマンの相関係数で0.414であり、表情認知課題時のみの眼球運動による予測結果をRMSEで0.92上回る結果となった。本稿ではその詳細についてまとめる。

2. 実験方法

本研究での調査項目は、表情認知課題時の眼球運動を用いる時のみと比べて、視点取り課題を行ったときの眼球運動を

同時に用いることによる、自閉傾向の予測精度の向上である。SRS-2を自閉傾向の評価尺度として用い、スコアの予測精度によって評価する。

2.1 実験参加者

本研究は奈良先端科学技術大学院大学の倫理承認を得て行われた。本研究では22から35歳の28名(男性11名, 女性17名)のデータを収集した。全実験参加者に対して、書面及び口頭で説明を行い参加の同意を得た。また、全実験参加者に対して、本研究の予測対象となる、SRS-2[6], Kikuchi's scale of social skills-18[7], 新版STAI状態-特性不安[8]を取得した。本研究では、SRS-2に関してのみ報告対象とする。また、本実験では、バーチャルエージェントの心の理論の程度による信頼性の評価[9], ソーシャルスキルトレーニングシステムのフィードバック評価に関するデータも取得しているが、本報告では対象としない。本研究により取得した眼球運動データなどは第2著者の田中に連絡の上、提供が可能である。

2.2 対人性応答尺度 (SRS-2)

SRS-2は、65問の質問を含む客観的な評価指標である。SRS-2は本来、自閉スペクトラム患者を評価するために設計されているが、様々な精神疾患を鑑別することも可能である。さらに、その有効性は疾患群だけでなく、健常者についても調べられている[6]。したがって、健常者に対しても自閉傾向を測ることができる指標となっている。SRS-2は

- Social Awareness
- Social Cognition
- Social Communication
- Social Motivation
- Restricted Interests and Repetitive Behavior

の5つの下位尺度によって自閉傾向を測っている。点数が高いほど自閉傾向が弱いとされる。本実験参加者28名のSRS-2の平均と分散は56.46と26.03であった。

2.3 使用した課題

本研究では、表情認知課題と視点取り課題を用いた。表情認知課題は成人版表情認知検査 (FEIT)[10], 視点取り課題はSamsonらのLevel-1 Visual Perspective Task (VPT)[11]を用いた。

2.3.1 FEIT

FEITの眼球運動取得時の実験の流れを図1に示す。答えを声に出してもらおう旨の教示を行った後、視線を中心に誘導するために、十字を1秒間提示した。その後、5秒経過するか、または表情がわかったタイミングでマウスでクリックすることによって7つの選択肢(喜び, 悲しみ, 恐怖, 怒り, 驚き, 嫌悪, 感情なし)が表示され、その中から選択してもらった、課題は全部で7つの表情に対して3枚ずつを無作為に出題し、21問行った。

2.3.2 VPT

次に、VPTの実験の流れを図2に示す。この課題と自閉群との関係は調査されており[12], 自己視点、もしくは他者視点で赤いドットが何個見えるかを確認し、教示と実際の画像の内容が合っているかを答える課題である。最初に750ミリ秒視線を中心に誘導するための十字が表示される。500msの

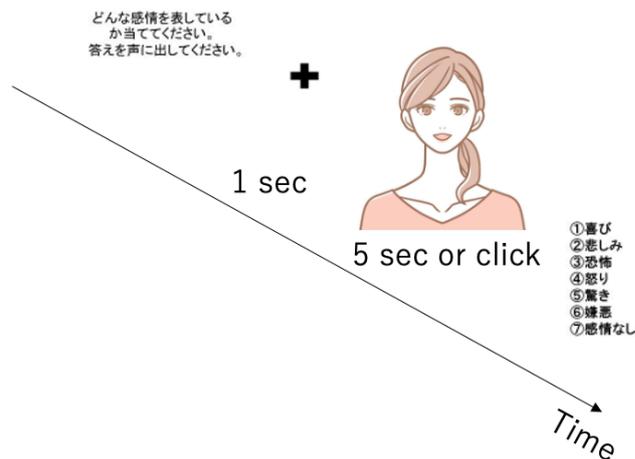


図1 表情認知課題の流れ。5秒間、またはクリック操作により選択肢が現れ、次の問題へと移る。全部で21問ある。また、著作権の関係により、ここで示している画像は実際に使った画像ではなく模式図となっている。

休止の後、「あなた」、もしくは「かれ」と表示され、500ミリ秒の休止の後、赤いドットの数を指定するための数字が表示される。最後に中心に人、左右の壁に赤いドットがある画像が表示され、2000ミリ秒経過するか、教示内容と状況が一致している場合 (Matching) は左クリック、一致していない場合 (Mismatching) は右クリックすることが求められた。自己視点 (Self), 他者視点 (Other) を答える場合と、視点間で見え方が一致している場合 (Consistent) としていない場合 (Inconsistent) での4条件が存在する。今回はこの課題を Self-Consistent 条件11課題, Other-Consistent 条件14課題, Self-Inconsistent 条件13課題, Other-Inconsistent 条件10課題, ドットが表示されないフィルター条件4課題の合わせて52課題を行った。課題は無作為に出題された。また、分析には Matching の場合のみを使うことが推奨されており[11], 本研究においても Matching の場合のみのデータを用いて分析を行った。

2.4 眼球運動と特徴量の取得手順

これらの課題を行っている際の眼球運動を取得した。使用した機器は Tobii Pro Fusion で、サンプリングレートは120Hzであった。課題を表示するために使用したディスプレイのサイズは1920×1080であった。眼球運動の分析には Tobii Pro Lab (Version 1.145) を使用した[13]。眼球運動を取得する際に Areas of Interest (AOI) を設定する必要があるが、本研究では、FEITの画像においては、目、口、顔に設定し、VPTにおいては中心に立っている人、赤いドットが表示される右側、左側の壁に設定した。取得した眼球特徴は Fixation の回数 (Number of fixation) と、Saccade の回数 (Number of Saccade) である。Fixation とは特定の位置を見つめるような眼球運動であり、Saccade とは、両方の目が同時に同じ方向に素早く動くような眼球運動のことである。FEITにおいては7つの表情ごとの平均を算出し、VPTにおいては、4条件で Number of Fixation と Number of Saccade を取得し、条件ごとに平均を算出した。また、課題に対する応

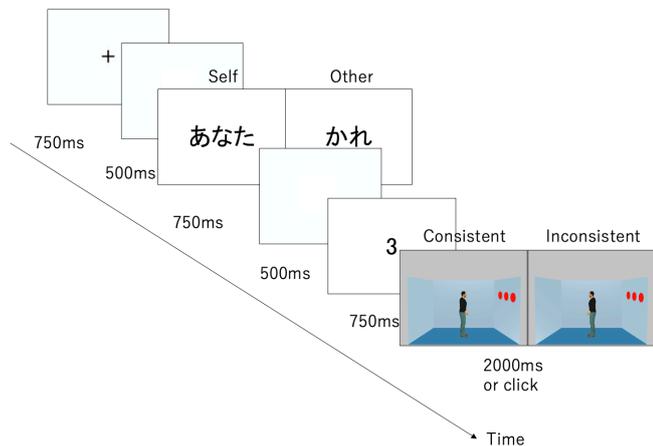


図2 VPTの流れ。「あなた」、もしくは「かれ」の教示の後、赤いドットの個数を指示するための数字が表示され、課題の画像が表示される。被験者は教示と最後に表示される画像の状況が合っている場合は左クリック、合っていない場合は右クリックをすることが求められた。

表1 各課題ごとに取得した特徴量の一覧。合計で55個の特徴量を取得した。

課題	特徴量	条件	合計特徴量	
FEIT	Number of Fixation at eyes	Happiness	35	
	Number of Fixation at mouth	Sadness		
	Number of Fixation at face	Fear		
	Number of Saccades	Anger		
	Response Time			Surprise
				Disgust
Neutral				
VPT	Number of Fixation at human	Self-Consistent	20	
	Number of Fixation at right wall	Self-Inconsistent		
	Number of Fixation at left wall	Other-Consistent		
	Number of Saccades	Other-Inconsistent		
	Response Time			

答速度 (Response Time) も取得し、合計55個の特徴量を取得した。抽出した特徴について整理したものを表1に示す。

2.5 予測手法

次に、予測手法について説明する。今回28名のデータを取得したが、そのうち1名のデータの眼球運動を取得できていなかった。そのため27名のデータを用いて機械学習モデルを構築した。サンプルサイズ ($n=27$) に対して特徴量 ($k=55$) が極端に多い場合、次元の呪いに対処する必要がある。そこで本研究では、相互情報量による次元削減を入力の前に行なった。相互情報量は一つの特徴量を x 、目的変数を y 、それぞれの確率を $p(x)$ 、 $p(y)$ 、同時確率を $p(x,y)$ として以下のように計算できる。

$$I(x,y) = \sum_{x,y} p(x,y) [\ln p(x,y) - \ln p(x)p(y)]$$

本研究では、過学習を防ぐために Training score と Cross validation score によって入力特徴の数を調整し、相互情報量が高い

表2 SRSを予測した各モデルの結果。各指標で最良の結果を太字にしている。

モデル	特徴セット	R^2	RMSE	ρ
PLS Regression	FEIT	0.053	25.32	0.400
	VPT	-1.11	37.83	-0.319
	FEIT+VPT	0.121	24.40	0.414
Linear Regression	FEIT	-0.009	26.13	0.321
	VPT	-2.56	49.11	-0.440
	FEIT+VPT	0.077	25.00	0.400

表3 PLS Regressionの上位5つの特徴量とその回帰係数。

特徴セット	特徴量	回帰係数
FEIT	Number of saccades Sadness	-13.17
	Number of saccades Surprise	-6.39
	Number of fixation at face Sadness	5.43
	Number of saccades Anger	-4.62
	Reaction time Surprise	3.68
FEIT+VPT	Number of saccades Sadness	-13.17
	Number of saccades Surprise	-6.19
	Number of fixation at face Sadness	5.41
	Number of saccades Self Inconsistent	4.85
	Number of fixation at human Self Consistent	-4.68

順に8つの特徴を入力することとした。また、今回は Linear Regression と Partial Least Square Regression (PLS) を用いてモデルを構築し、パラメータ調整と評価を行うために Nested Leave one out cross validation を行った。入力特徴は標準化を行っている。モデルに入力する特徴をFEITのみの場合、FEITとVPTの特徴を同時に入力した場合の2条件で比較した。出力はSRSスコアである。評価指標には決定係数 R^2 、RMSE、Spearmanの相関係数 ρ を用いる。また、実装には Scikit-learn を用いた [14]。

3. 結果

PLS Regression, Linear Regression を用いて予測した結果を表2に示す。PLSにFEIT, VPTの眼球特徴を両方用いた場合のモデルの評価指標が決定係数0.121, RMSE24.40, 相関係数0.414と最も高い値を示した。今回最良の相関を得た PLS Regression に、全ての特徴を用いた場合のモデルの予測値と真値の相関係数 ρ に対して、無相関検定を行った結果、 $p=0.0317$ となり $p<0.05$ の水準で有意な結果が得られた。また、FEITのみの特徴、すべての特徴を用いてSRSを予測した際の真値と予測値の散布図を図3, 4に示す。また、PLS Regressionは線形モデルのため、回帰式の係数を算出することができ、Leave one out cross validation によって得られた、各モデルの回帰係数を平均した値の絶対値が大きかった特徴量を表3に示す。

4. 考察

今回の実験では、FEITのみの特徴を用いた場合に比べ、FEITとVPTの眼球運動を用いた場合のモデルがいずれの評価指標においても上回る結果となった。また、VPTのみの眼球運動を用いた場合は予測することができなかった。これは、表3の

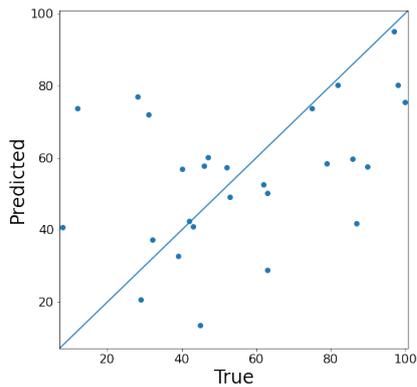


図3 FEIT のみの特徴を用いて SRS を予測した際の真値と予測値の散布図。真値と予測値が近いほど青い線に近づく。

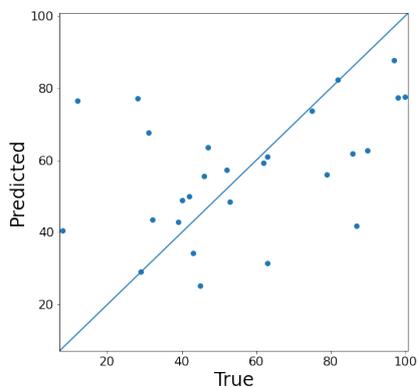


図4 すべての特徴を用いて SRS を予測した際の真値と予測値の散布図。真値と予測値が近いほど青い線に近づく。

回帰係数を見ると、FEIT から得られる特徴に対して、特に大きい回帰係数が乗っているからだ考える。上位3つの予測に寄与している特徴はFEITのものであり、FEIT のみの特徴を使用した場合のモデルと変わらないが、下位の2つに関しては、FEIT のものではなく、VPT のものが影響している。このことから、SRS 予測に対して、大きな影響を与えているものはFEIT の特徴であり、VPT のものはFEIT に比べて影響が小さいことがわかる。また、自己視点を問われた際の Saccade 回数には正の係数が、自己視点を問われた際の Fixation 回数には負の係数がかかっている。これは、自己視点を問われた際に、Saccade 回数が増えると SRS スコアが大きくなり、Fixation 回数が増えると SRS スコアが低く予測されることを意味する。この課題を解く際、健常群では自己視点を問われていても無意識的に他者視点からのドットの数を数えてしまう、altercentric effect が起こることが知られている [11], [15]。このことから、SRS スコアが高い人、つまり自閉傾向が高い人ほど、altercentric effect が起きづらいことが示唆される。

5. ま と め

本論文では表情認知課題に加え、VPT 時の眼球運動を用い

ることによって、自閉傾向の予測精度を改善できることを示した。このことから、異なる認知活動を測る課題によって、複合的に自閉傾向を評価することができ、結果として予測精度が改善されたことが示唆される。また、VPT を追加した場合としていない場合での回帰係数の値を比較したところ、上位3つは同じであったが、下位二つはVPT 課題のものであった。この結果から、SRS スコアが高い人は、altercentric effect が起きにくいことが示唆された。今後の課題として、サンプルサイズが27と非常に小さいことが挙げられる。より多くのデータを集めての検証が必要である。また、課題が増えることによる被験者への影響も考慮する必要がある。FEIT, VPT ともに出題数による影響を調査する必要がある。また、今回VPT によって、自閉傾向のさまざまな側面のうち、どの側面をFEIT のみの時と比べて評価できるようになったのかを詳細に分析していく必要がある。

謝 辞

本研究はCREST(グラント番号: JPMJCR19A5)の支援によって行われた。

文 献

- [1] K. Ochi, N. Ono, K. Owada, M. Kojima, M. Kuroda, S. Sagayama et al. "Quantification of speech and synchrony in the conversation of adults with autism spectrum disorder.", PLOS ONE, 2019.
- [2] K. Iwauchi, H. Tanaka, K. Okaazaki, H. Iwasaka, Y. Matsuda, M. Uratani, T. Morimoto, S. Nakamura, "Detecting of Psychiatric Disorders from Eye Movement Considering the Types of Emotional and Cognitive Stimuli.", HCG Symposium, 2021.
- [3] M. Jiang et al., "Predicting Core Characteristics of ASD Through Facial Emotion Recognition and Eye Tracking in Youth.", 2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC), 2020.
- [4] S. Baron-Cohen, "Theory of mind and autism: A review.", International review of research in mental retardation, vol.23, pp.169 - 184, 2000.
- [5] P. Michelon and J.M. Zacks, "Two kinds of visual perspective taking.", Perception Psychophysics 68, pp.327 - 337, 2006.
- [6] John N. Constantino, "Social Responsiveness Scale, Second Edition (SRS-2).", Back. Western Psychological Services. 2012.
- [7] K. Kikuchi, "Shakaiteki-sukiru-wo-hakaru-kiss-18 handbook.", The social skills are measured, The handbook of Kiss-18, Tokyo, Japan, Kawashima, 2007.
- [8] C.D. Spielberger, "State-trait anxiety inventory for adults.", 1983.
- [9] H. Tanaka, T. Saga, K. Iwauchi, and S. Nakamura, "Trustworthiness of Virtual Agents in terms of Theory of Mind Levels and Social Skills Training.", HIP, 2022
- [10] 成人版表情認知検査, <http://www.saccess55.co.jp/untitled406.html>
- [11] D. Samson, IA. Apperly, JJ. Braithwaite, BJ. Andrews, SE. Bodley Scott. "Seeing it their way: evidence for rapid and involuntary computation of what other people see.", J Exp Psychol Hum Percept Perform. 2010 Oct;36(5):1255-66. doi: 10.1037/a0018729. PMID: 20731512.
- [12] H. Doi, C. Kanai, N. Tsumura, K. Shinohara, and N. Kato, "Lack of implicit visual perspective taking in adult males with autism spectrum disorders.", Res Dev Disabil, 2020.
- [13] Tobii Pro AB. Tobii pro lab (Version 1.145), Computer software, 2014.
- [14] P. Fabian et al., "Scikit-Learn: Machine Learning in Python", Journal of Machine Learning Research, 2011.
- [15] H.J. Ferguson, I. Apperly, and J.E. Cane. "Eye tracking reveals the cost of switching between self and other perspectives in a visual perspective-taking task.", Q J Exp Psychol (Hove), pp. 1646-1660, 2017.