

# AIoT行動変容学会 第10回研究会論文集

(Proceedings of the Academy of Behavior Transformation by AIoT 10th Research Meeting)

2025年12月8日(月)

慶應義塾大学三田キャンパス

# 目 次

- プログラム … p. 1
- 一般講演予稿集 … p. 2~49
- ポスター・デモ・企業発表予稿集 … p. 50~65

## BTI10

2025年12月8日(月) @慶應義塾大学 三田キャンパス

## — プログラム —

12月8日（金）@G-lab

10:30	受付開始	
11:00	オープニング	
11:10		
セッション1	MBA : Meal Balance Adviser -食事バランスを改善させる弁当トレイ- 脈拍変動を用いたデスクワークにおける休憩方法による集中回復への効果検証	比留間 瑞穂（青山学院大学） 羽持 涼花（青山学院大学）
12:10	休憩	
13:10		
セッション2	行動変容における両面性を乗り越えるための動機づけメッセージ生成手法に関する選好調査 過去施策の有効性推定に基づくユーザー特性と行動を考慮した PUSH 配信手法 Big Five性格特性に基づく質問を用いた内省誘発と決定遅延評価の試み	有賀 玲子（NTT） 三浦 優明（株式会社NTTドコモ） 上野 貴弘（福山大学）
14:40	休憩	
15:00		
ポスター/デモセッション	AIとのコミュニケーションによる環境配慮行動の誘導可能性－ごみ処理ゲームを題材とした検証－ 運動予測可能性への選好に基づくASDリスク発見の可能性 一般運転者の避讓行動促進に向けたVR緊急走行体験システムの開発 Human-AI TeamingにおけるAIの役割越境行動と信頼形成 可搬型IoTを活用した多人数対象昇降・転倒初動検知に向けたIMUデータ準備時間短縮システム 上半身姿勢情報を用いた作業専念度のリアルタイム推定モデルとアバタの表情表示連携 複数のAIoTアバタを活用した情報提供システムの構築	日室聰仁（NECソリューションイノベータ株式会社） 大森幹真（早稲田大学人間科学学術院） 西浦裕生（兵庫県立大学） 武安大樹（東京科学大学） 小島涼太郎（東京電機大学） 山口美佳（東京電機大学） 木田 勇輝（慶應義塾大学大学院）
16:30	休憩	
16:45		
企業セッション	ドライバーの運転行動変容による自動車からの排出ガス低減に関する研究	池本 雅里（トヨタ自動車株式会社）
17:15	クロージング	
17:30	解散	
18:00	懇親会 @肴亭 田町店	

# MBA : Meal Balance Adviser

## -食事バランスを改善させる弁当トレイ-

谷塙 碧 <sup>†1</sup>

青山学院大学

比留間 瑞穂 <sup>†2</sup>

青山学院大学

ロペズ ギヨーム <sup>†3</sup>

青山学院大学

### 1. はじめに

近年、日本ではバランスの取れた食事を見る頻度が減少している。世代別の調査では、特に若い世代において、過半数が主食・主菜・副菜を3つそろえて食べることが週3日以下である[1]。食事をする際、野菜から食べると野菜に含まれる食物繊維が糖質・脂質・コレステロールの消化吸収を遅らせ、食後の血糖値上昇を抑制することが知られており、食べる順番が健康に影響することが明らかになっている[2][3]。主食・主菜・副菜のバランスを意識した食事の頻度が少ないほど、恒常に不足している栄養素の数が多い傾向が認められている[4]。

食事バランスに対する意識を高めることで、食事行動を見直すことにつながることが期待できるため、本研究では、新しい手法として、お弁当トレイ型デバイスの周囲を発光させることでバランスよく食事することを視覚的に促すデバイスを仕組み (MBA: Meal Balance Adviser) を提案・開発し、その効果を検証した。

### 2. 関連研究

Kimらは、光と振動による食事速度ガイドシステム「Slowee」を提案している[5]。本システムは、ヘッドホン型EMGセンサとネックレス型ピエゾセンサで咀嚼・嚥下を検出し、食事速度に応じてライトとリストバンド型デバイスで光と振動のフィードバックを行う。比較実験では、リアルタイムフィードバックにより咀嚼回数増加と食事ペース改善が確認された。

門村らは、食事行動検出と健康的な食習慣を促進するためのデバイス「Sensing Fork」と、それに連動するモバイルゲームシステムを提案している[6]。フォーク型デバイスに加速度・ジャイロセンサを搭載し、食事動作を検出してBluetooth経由でモバイルアプリへ送信、ゲーム形式でフィードバックを提供する。比較実験では、フィードバッ

クにより咀嚼回数増加や食事時間延長が確認され、早食い防止にも有効であることが認められた。

小原らは、視覚的フィードバックで早食いや偏食を改善するシステムを提案している[7]。骨伝導マイクで咀嚼回数や発話時間を測定し、測定結果に応じてLED照明の色変化や点滅運動を制御する。評価実験では、緑や青の照明と1秒間隔の点滅が咀嚼回数増加に効果的であり、発話抑制にも効果があった。一方で、点滅が速すぎる場合は不快感が生じ、快適性と効果のバランスが重要であることが示された。

藍場らの研究では、野菜を先に摂取する場合や、三角食べのようにバランス良く摂取する場合では血糖値の上昇が抑制されることが確認された[8]。また、野菜やたんぱく質の摂取がインスリン分泌を増加させ、糖質代謝を促進する可能性が示された。結果から、食事順序の工夫が食後血糖値を効果的にコントロールするために有効な方法であることが示された。

村上らの調査では、白飯を連続して摂取するような「偏食型」の食べ方が、栄養の偏りや満足度の低下につながる可能性があることが認められた[9]。一方で、主食・主菜・副菜をバランスよく摂取する「三角食べ」では、栄養バランスの向上や満腹感の増加に効果的であることが確認された。

松尾らの研究では、米飯・肉料理・野菜料理の3つの食材を用い、それらの摂取順序を変化させた場合の食後血糖値の変動を比較した[2]。野菜を最初に食べる場合が最も食後血糖値の上昇が緩やかで、次いで「三角食べ」であった。米飯を最初に食べる場合は、食後血糖値の上昇が急激で、ピーク値も高いことが確認された。

小林らは、照明条件による食品に対する食欲についての調査を行った[10]。光色は、赤・緑・青をそれぞれ濃・淡で分けた6種類とし、食品は20品目で比較した。被験者は「抵抗なく食べられる」「抵抗はあるが食べられる」「食べられない」の選択肢が与えられた。淡色光では「抵抗なく食べられる」という回答が75%を超えるものが多かったのに対し、濃色光では割合が低くなり、50%を下回るものもあった。

同じく小林らは、カフェ環境における光色とBGMの違いが会話行動および心理評価に与える影響を調査した[11]。実験では、テーブル面照度50lxの下で白・青・赤・緑の4

Conference Manuscript Format for Academy of Behavior Transformation by AIoT (BTI)

<sup>†1</sup> AOI TANINO, Aoyamagakuin University

<sup>†2</sup> MIZUHO HIRUMA, Aoyamagakuin University

<sup>†3</sup> GUILLAUME LOPEZ, Aoyamagakuin University

種類の光色と、等価騒音レベル 65 dB の 4 種類の BGM (ハウス, ボサノバ, バラード, ヒーリング) を組み合わせた 16 条件を設定し、同性の友人同士による会話を観察した。結果として、光色単体による行動変化はほとんど見られなかったが、話しかけやすさや視線に関する心理評価は光色に依存し、男性は赤色光、女性は緑色光で評価が低下した。また、光色と BGM の組み合わせによる複合的な影響も確認された。しかし、本研究は光色と BGM の組み合わせに焦点を当てており、光色単独の影響に関する考察は限定的である。

豊島らの研究では、食卓・食器・食物の配色が嗜好に与える影響を検討した [12]。食卓 8 色、食器 8 色、食物 8 色の計 512 通りの組み合わせについて「美味しそうに見えるか」を順位付けした結果、上位は白・オレンジ・黄色など彩度が高く明るい暖色系であり、赤も比較的多く含まれた。一方、下位は明度・彩度ともに低い暗色系であった。食卓・食器・食物別の比較では、いずれも最高評価は白、最低評価は黒であり、特に食物において評価の幅が大きく、暖色系が高評価、寒色系が低評価となる傾向が示された。

篠田らは、視覚や聴覚などの外部刺激が人間の行動選択や誘導に与える影響を検証した [13]。特に、光刺激による行動誘導の可能性を探究した。視線計測装置を用いた展示空間での実験では、被験者が視界内の明るい部分に視線を向けやすい傾向が確認され、光刺激が短期的な行動誘導に有効である可能性が示された。また、T 字路や迷路を用いた実験では、光源が行動選択に一定の効果を持つことが示されたが、その効果は光の強度、配置、明暗差の強調具合に依存することが明らかとなった。

### 3. システム概要：MBA の仕組み

#### 3.1. MBA の構成

##### 3.1.1 MBA の概要

MBA システムは、弁当トレイ型デバイスであり、感圧センサを用いて摂食動作を検出し、トレイ周囲に配置した LED テープで次に食べるべきお弁当の領域を提示することで、三角食べを促進し、食事バランスの改善を支援する。対象は、偏食を避けて適切な食習慣を身に付けたい人や、早食いや過食の改善を目指す人であり、偏った食事をしがちな場面や、子どもが食事バランスの取れた食べ方を学ぶ際の利用を想定している。利用イメージを図 1 に示す。

##### 3.1.2 ハードウェア構成

MBA は、M5Stack-Gray, RGB-LED テープ、およびトレイ裏面に配置した 4 つの感圧センサを主要構成要素とする弁当トレイ型デバイスである（図 2）。トレイのサイズは、実験で使用する弁当（約 20cm × 16cm）を収容可能なサイ

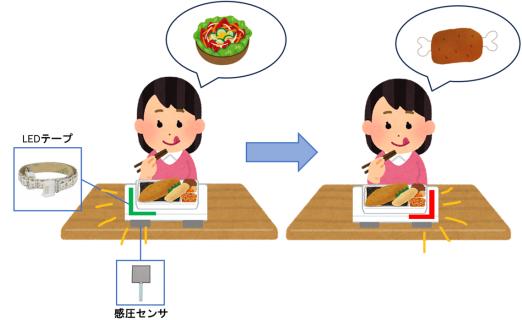


図 1 MBA システムの利用イメージ

ズ（24.5cm × 22cm）で設計した。弁当は 4 区画に分割されており、各区間の中央に感圧センサを配置し、摂食時の圧力を検出する。

トレイの四隅にはレゴブロックを 2 個ずつ設置し、弁当の位置を安定させると共に、RGB-LED テープの取り付けが可能な構造とした。また、感圧センサの上にも小型のレゴブロックを配置することで、圧力変化が検知しにくくなる問題を回避し、箸による摂食動作時の圧力を効果的に検出できるようにした。

M5Stack-Gray と RGB-LED テープは Grove 互換ケーブルで接続し、感圧センサで取得した圧力値に基づいて M5Stack-Gray が RGB-LED テープの照明を制御する。RGB-LED テープには M5Stack 社製 NeoPixel 互換 LED テープ（200cm、型番 A035-E）を採用し、感圧センサには Alpha 社製 MF02-N-211-A01（図 3）を使用した。

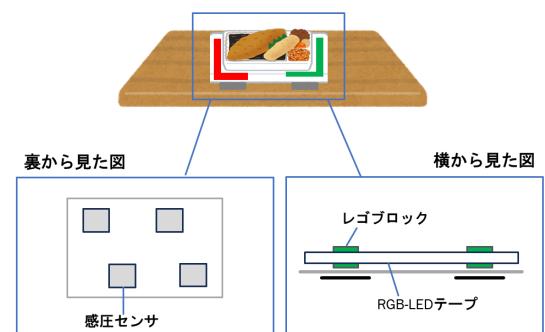


図 2 MBA システムの構成

##### 3.1.3 ソフトウェア構成

感圧センサは 30 ミリ秒ごとに圧力測定を行う。M5Stack-Gray は、感圧センサで測定した圧力値に基づき、弁当のどの区画で摂食動作が行われたかを判定し、対応する領域の LED テープを点灯させることで視覚的フィードバックを提供する。



図 3 使用した各デバイス

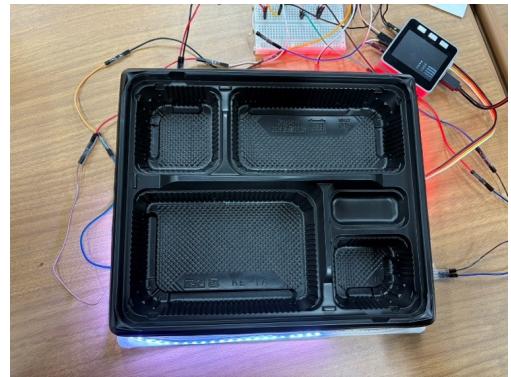


図 5 MBA デバイスでの食事終了時風景

### 3.2. 各機能の操作基準

#### 3.2.1 LED 点灯区域決定の仕組み

図 4 に、本実験で使用する弁当の区域を示す。それぞれの区画の下に感圧センサを配置し、摂食動作を検出すると、別の区画の LED テープが点灯する仕組みとなっている。図 5 は、デバイス使用時に区画ごとに LED テープが点灯している様子を示す。また、図 6 には、複数の点灯パターンを示す。点灯の条件は以下の通りである。

1. 主食を摂取した場合：主菜と副菜の区域が点灯
2. 主菜を摂取した場合：主食と副菜の区域が点灯
3. 副菜を摂取した場合：主食と主菜の区域が点灯



図 4 弁当の区域

#### 3.2.2 摂食動作の判定

摂食動作を正確に判定するため、感圧センサの圧力出力値が特定の閾値を超えた回数に基づき閾値を設定した。同一の弁当区画で摂食動作を行ってもらい、その際の圧力変化を被験者 3 名に対して 5 回ずつ、計 15 回測定した。圧力センサの出力値を 0~4000 (単位のない AD 変換値) まで 500 刻みで区切り、それぞれの閾値を超えた反応回数をカウントし、最も反応回数が多い値を閾値として採用した。

結果として、主菜は圧力値が 2000 を超えた場合に一口食べられたと判定し、副菜は 1500 を超えた場合に判定するこ

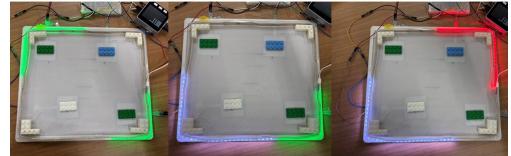


図 6 LED テープ点灯風景

とした。一方、主食はシステム開始時点で圧力値が高いことが確認されたため、被験者 9 名を対象に、開始後 10 秒間の平均圧力を測定した結果、平均値は約 2800 であったため、この範囲で最も反応回数が多かった 3000 を閾値として設定した。

#### 3.2.3 適切な咀嚼時間の判定

適切な咀嚼時間を設定するため、咀嚼ランク表 [14] を参考にした。ポテトサラダはランク 2 (咀嚼回数 20~30 回)、ハンバーグや唐揚げ、アジフライはランク 3 (咀嚼回数 30~40 回) に分類される。そこで、20 回および 30 回の咀嚼に要する時間を被験者 3 名で各 5 回、計 15 回測定した (表 1, 表 2,)。結果、20 回の咀嚼には最低 9 秒、30 回には最低 14 秒を要した。以上を踏まえ、副菜摂取後は 10 秒、主食・主菜摂取後は 15 秒経過すると LED テープが点灯するよう設定した。

表 1 20 回の咀嚼にかかった時間

被験者 1	被験者 2	被験者 3
11.60	12.23	9.67
11.58	9.93	9.33
11.98	9.72	9.63
11.23	10.08	9.60
11.83	9.92	9.85

表 2 30 回の咀嚼にかかった時間

被験者 1	被験者 2	被験者 3
16.84	17.84	14.66
16.93	15.34	14.49
17.89	14.91	14.54
16.86	15.82	14.18
17.36	15.14	14.38

### 3.2.4 RGB-LED テープ制御における色分けの基準

RGB-LED テープの色設定は、食品の種類に対して視覚的に連想しやすく、味覚認識に影響を与えない色を選定した [15]。さらに、美味しそうに見える色として白・オレンジ・黄色、次いで赤が好まれるという報告を参考に、主食を白、主菜を赤、副菜を緑に設定した。また、光の彩度が強すぎると食欲が低下する可能性があるため、輝度は 30 と低めに設定した。

## 4. 有用性評価実験

### 4.1. 実験概要

本実験では、食事バランスの向上および三角食べの促進効果を確認することを目的としている。開発システムが被験者の食事行動に与える影響を評価し、実験後のアンケートを通して、システム使用時の負荷や満足度、実用性についても検証を行う。

### 4.2. 評価手法

開発システムの有無による効果を検証するために、食事開始から終了までの総食事時間から会話や飲料摂取の時間を除いた総咀嚼時間を測定し、弁当 4 区分における連続摂食の減少を観察した。さらに、システムが提示する光の指示に対する反応を評価するため、指示に従った回数と従わなかった回数の割合を算出し、三角食べ促進効果を確認した。

加えて、システム使用時の負荷を評価するため NASA-TLX[16] による作業負荷の量定化を行い、満足度と実用性の確認には SUS (System Usability Scale) [17] に基づくアンケートを実施した。

### 4.3. 実験手順

本実験では、20 代の男女 8 名を被験者とし、実験の様子をスマートフォンで記録した。被験者には 2 日間に分けて、開発システムを使用する場合と使用しない場合で食事を行ってもらった。食事には箸を用いて弁当を摂取し、被験者の好みに合わせて 3 種類の主菜（ハンバーグ、唐揚げ、アジフライ）を用意した。弁当のサイズはシステムに適合する中

サイズで統一し、各被験者は両条件で同じ種類・サイズの弁当を食べた。実験終了後、システム使用時の負荷評価には NASA-TLX を、実用性評価には SUS を用いたアンケートを実施した。実験の様子を図 7 に示す。

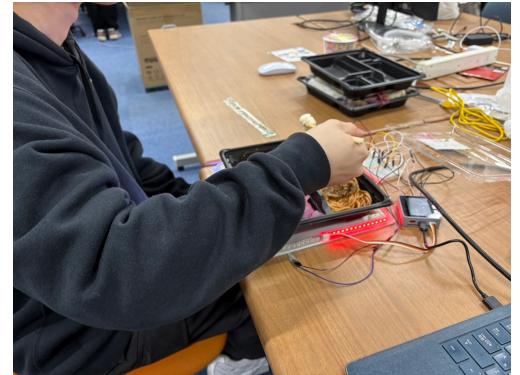


図 7 実験の様子

## 5. 結果

### 5.1. 実験結果

#### 5.1.1 MBA の有無による総食事時間の比較結果

MBA の有無による総食事時間の比較結果を表 3、表 4、および図 8 に示す。表 3 は各被験者の食事時間と増加率、表 4 および図 8 は平均食事時間と増加率を示している。結果、被験者 8 名全員で総食事時間が増加し、平均で 7 分 38 秒増加した。さらに、各総食事時間に対して t 検定を実施したところ、 $p < 0.05$  の水準で有意差が認められた。

表 3 開発システムの有無による総食事時間（分：秒）と増加率（%）

被験者番号	開発システムなし	開発システムあり	増加率（%）
1	9:09	14:53	63
2	12:09	22:19	84
3	13:39	17:48	30
4	19:06	32:06	68
5	9:58	18:09	82
6	20:16	32:49	62
7	16:21	19:31	19
8	15:32	19:41	27

表 4 開発システムの有無による総食事時間（分：秒）の平均

開発システムなし	開発システムあり	増加率（%）
平均	14:31	22:09

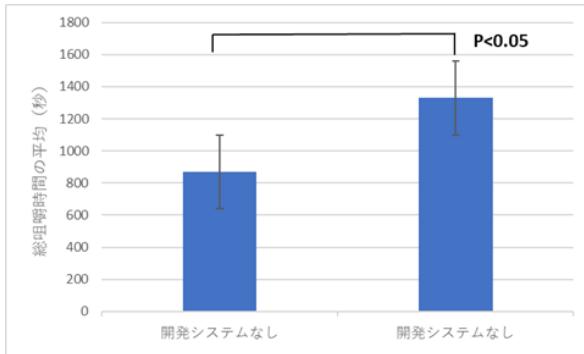


図 8 開発システムの有無による総咀嚼時間の平均 (秒)

### 5.1.2 MBA の有無による総咀嚼時間の比較結果

本実験では、会話や飲料摂取を制限せず、食べ物を口に入れて咀嚼している時間の合計を「総咀嚼時間」と定義し計測した。MBA 使用時と未使用時の総咀嚼時間の増加率を表 5 に、平均総咀嚼時間と増加率を表 6 および図 9 に示す。結果、被験者 8 名全員で総咀嚼時間が増加し、平均で 5 分 19 秒延長した。さらに、各総咀嚼時間に対して t 検定を実施したところ、 $p < 0.05$  の水準で有意差が認められた。

表 5 開発システムの有無による食事時間と増加率

被験者番号	開発システムなし	開発システムあり	増加率 (%)
1	7:11	13:21	86
2	9:35	18:32	93
3	9:43	15:35	60
4	10:32	18:50	79
5	6:10	9:57	61
6	11:17	16:30	46
7	8:20	10:22	24
8	8:56	11:10	25

表 6 開発システムの有無による実際の食事時間（分：秒）の平均

	開発システムなし	開発システムあり	増加率 (%)
平均	8:58	14:17	59

### 5.1.3 MBA の有無による連続摂食回数の比較結果

本実験では、連続摂食行動を「4 区画に分けられた弁当から特定の区画を食べ、再び同じ区画から摂取する行動」と定義し、計測を行った。MBA 使用時と未使用時の連続摂食回数・減少回数・減少率を表 7 に示し、平均値を表 8 および図 10 に示す。結果、被験者 8 名全員で連続摂食回数が減少し、平均で 10 回減少した。さらに、各連続摂食回数に対して t 検定を実施したところ、 $p < 0.05$  の水準で有意差が認められた。

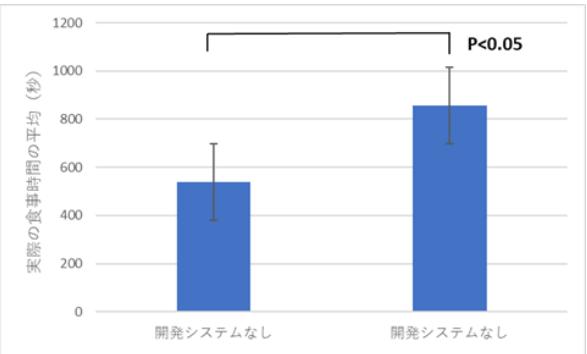


図 9 開発システムの有無による実際の食事時間の平均 (秒)

表 7 開発システムの有無による連続摂食回数

被験者番号	開発システムなし	開発システムあり	減少回数	減少率 (%)
1	18	6	12	67
2	25	5	20	80
3	10	6	4	40
4	16	8	8	50
5	13	5	8	62
6	9	3	6	67
7	10	7	3	30
8	21	6	15	71

表 8 開発システムの有無による連続摂食回数の平均

	開発システムなし	開発システムあり	減少回数	減少率 (%)
平均	15	6	10	58

## 5.2. アンケート評価の結果

### 5.2.1 NASA-TLX による MBA 利用における作業負荷評価の結果

実験後、図 4-2 に示したアンケートを基に、開発システムの有無による重み付け平均値を算出した（図 11、図 12）。結果、被験者 8 名全員で MBA 使用時の負荷が高く、平均値でも同様の傾向が確認された。各重み付け平均値に対して t 検定を実施したところ、 $p < 0.05$  で有意差が認められた。

### 5.2.2 SUS による MBA 利用におけるユーザビリティ評価の結果

実験後、MBA のユーザビリティを評価するため SUS アンケート（図??）を実施した。被験者ごとの回答結果を表 9 に示す。SUS スコアは各質問の得点を合計し、2.5 倍して算出した。結果、平均スコアは 83.8 点であり、最高評価は Q7 「たいていの人は、システムの利用方法をすぐに理解すると思う」で 4.9 点だった。一方、最低評価は Q8 「システムはとても操作しづらいと感じた」で 1.8 点であり、標準偏差も大きかった。

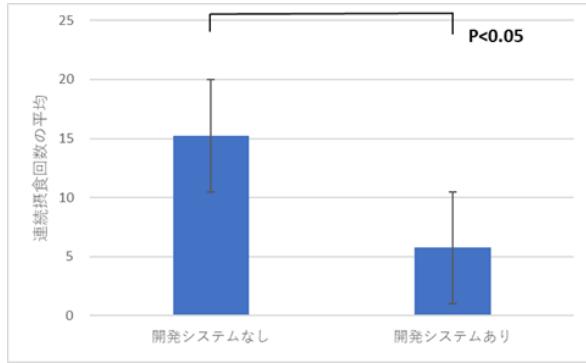


図 10 開発システムの有無による実際の食事時間の平均(秒)

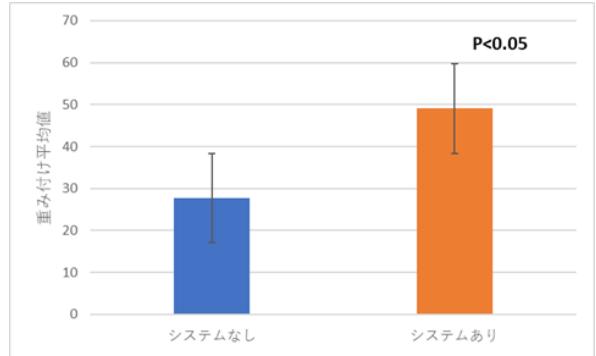


図 12 被験者の平均で算出した重み付け平均値

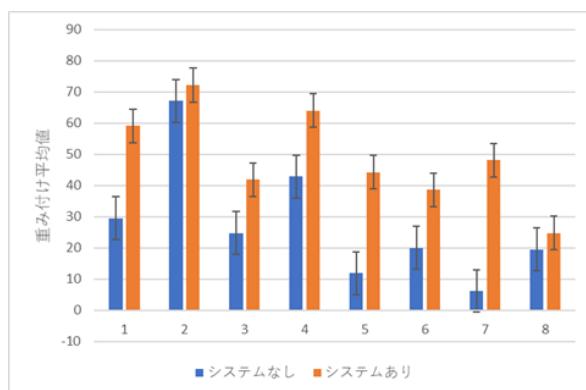


図 11 開発システムの有無による各被験者の重み付け平均値

表 9 SUS のアンケート結果

被験者番号	No.1	No.2	No.3	No.4	No.5	No.6	No.7	No.8	No.9	No.10	スコア
1	4	3	4	2	4	2	4	2	4	2	72.5
2	4	3	4	2	4	2	5	3	5	2	75
3	4	1	5	1	3	3	5	1	5	1	87.5
4	5	1	5	1	5	1	5	1	5	2	97.5
5	5	1	5	1	5	1	5	1	5	2	97.5
6	4	2	4	1	5	2	5	1	5	2	87.5
7	4	2	4	1	5	2	5	2	4	4	77.5
8	4	3	4	1	5	2	5	3	5	4	75
平均値	4.3	2	4.4	1.3	4.5	1.9	4.9	1.8	4.8	2.4	83.8
標準偏差	0.46	0.93	0.52	0.46	0.76	0.64	0.35	0.89	0.46	1.06	10.18

### 5.3. 考察

#### 5.3.1 総食事時間・総咀嚼時間・連続摂食回数の比較結果に関する考察

MBA を使用した場合、総食事時間は平均 7 分 38 秒、総咀嚼時間は平均 5 分 19 秒延長し、いずれも有意差が認められた。また、連続摂食回数は平均 10 回減少し、三角食べの促進効果が示唆された。これらの結果から、光の色分けと持続点灯による視覚的フィードバックが食事ペースの調整に有効であると考えられる。一方、一部の被験者ではシステム指示を無視する行動が確認され、その要因として咀嚼

推奨時間の設定が長すぎたことが考えられる。食材が口内からなくなり、複数食材を同時に食べたいという意識から、LED 消灯前に別区画を摂食する行動が発生したと推測される。今後は、咀嚼推奨時間の見直しや「口内調味」を考慮したシステムの提案が必要である。

#### 5.3.2 NASA-TLX と SUS の結果に関する考察

NASA-TLX 評価では、全被験者で MBA 使用時の負荷が増加し、有意差が認められた。これは、システム操作や指示理解に伴う認知負荷が影響した可能性がある。ただし、自由記述では「バランス良く食事できる」「満腹感が増した」「野菜から食べようと思った」など肯定的な意見が多く、MBA が食事習慣改善に寄与することが示唆された。一方、「箸がしなって食べにくい」という指摘もあったため、圧力検出のために設置したレゴブロックの配置改善が必要である。

SUS 評価では平均スコア 83.8 点と標準ガイドラインの 68 点を大きく上回り、A ランクに相当した(表 10)。標準偏差を考慮した 73.62 点も高水準であり、MBA は食事に支障が少なく、満足度の高いシステムであると考えられる。一方、Q1「システムをしばしば利用したい」と Q9「利用できる自信がある」は評価のばらつきが大きく、低めであった。これは、次に食べるべき区画を LED 点灯で示す仕組みや、摂食動作を検知した後も LED が点灯し続ける理由(咀嚼推奨時間の提示)について説明が不足していたことが原因と考えられる。例えば、主食の区画が点灯し、摂食後も推奨咀嚼時間の間は点灯が続く仕様であるが、被験者は「検知されていないのではないか」と不安を感じた可能性がある。改善策として、LED 点灯に加え、液晶に経過時間を表示するなど、視覚的フィードバックを強化することで理解を容易にできると考えられる。

表 10 SUS スコアにおけるガイドライン

SUS Score	Grade	Adjective Rating
>80	A	Excellent
71 - 80	B	Good
62 - 71	C	Okay
51 - 62	D	Poor
<51	F	Awful

## 6. 結論

### 6.1. 本研究のまとめ

本研究では、LED テープによる視覚的フィードバックを用いて食事バランスの偏りや早食いの改善を目的として食事バランス推進システム「MBA : Meal Balance Adviser」を開発した。評価では、連続摂食回数・総食事時間・総咀嚼時間を比較し、SUS によるユーザビリティ評価と NASA-TLX による作業負荷評価を実施した。結果、MBA 使用時には総食事時間と総咀嚼時間が有意に延長し ( $p < 0.05$ )、連続摂食回数は有意に減少した ( $p < 0.05$ )。また、NASA-TLX では負荷が増加し、SUS スコアは平均 83.8 点と高評価であった。これらの結果から、MBA は食事ペースの調整と三角食べの促進に有効であり、食事習慣改善に寄与する可能性が示された。

本研究の取り組みに対する結論は以下の通りである。

1. NASA-TLX 評価より、視覚的フィードバックは摂食行動に負荷を与える、食事ペースを調整する効果がある。
2. 連続摂食回数の減少から、MBA は食事バランスを意識させ、三角食べを促進できる。

### 6.2. 今後の展望

本研究の実験は短期間で、被験者がシステムを使用したのは 1 回のみであった。今後は、長期的な日常利用によって食事バランス改善や健康的な食習慣形成に寄与するかを検証する追加実験が必要である。また、被験者からは「光を追いたくなる仕掛けが欲しい」や「レゴブロック上では弁当が不安定」という意見があり、視覚的誘導効果と安定性の向上が課題である。

さらに、現行システムは特定サイズの弁当に依存しているため、異なる形状やサイズに対応できる柔軟な設計が求められる。加えて、実験対象は 20 代男女 8 名のみであり、年齢差による影響の検証も必要である。咀嚼推奨時間は考慮したが、具材の量は考慮されておらず、副菜など量が少ない区画では食べ終わっても LED が点灯し続け、混乱を招いた。今後は、摂取量を考慮した判定機能を導入し、より精度と実用性の高いフィードバックを提供することが求めら

れる。

## 参考文献

- [1] 農林水産省：ゆっくり食べる (2022). (参照日 2024/1/22).
- [2] 今井佐恵子：野菜から食べる「食べる順番」の効果、大阪府栄養士会.
- [3] 稲井 玲子松尾 達博：米飯、肉および野菜の摂食順が若年者の食後血糖値に及ぼす影響、香川大学農学部学術報告、Vol. 75, No. 128, pp. 73–78.
- [4] Aiko Narumi-Hyakutake, Kimika Yamada, Y. Y.: Relationship between Frequency of Meals Comprising Staple Grain, Main, and Side Dishes and Nutritional Adequacy in Japanese Adults, *Nutrients*, Vol. 16(11), No. 1628 (2024 May 26).
- [5] Kim, J., Lee, K.-J., Lee, M., Lee, N., Bae, B.-C., Lee, G., Cho, J., Shim, Y. M. and Cho, J.-D.: Slowee: A smart eating-speed guide system with light and vibration feedback, *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pp. 2563–2569 (2016).
- [6] Kadomura, A., Li, C.-Y., Chen, Y.-C., Tsukada, K., Siiio, I. and Chu, H.-h.: Sensing fork: Eating behavior detection utensil and mobile persuasive game, *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, pp. 1551–1556 (2013).
- [7] 小原文、ギヨームロベズ：食事中における照明制御を活用した咀嚼発話意識向上に関する研究、研究報告ヒューマンコンピュータインタラクション (HCI) (2018).
- [8] 藍場 元弘・川東 美菜・河野 友晴・戎谷友希・藤本 侑希・橋田誠一：食べる順番による血糖値および尿中インスリン量の変動に関する研究 (2016). [https://www.maff.go.jp/j/syokuiku/minna\\_navi/topics/topics4\\_02.html](https://www.maff.go.jp/j/syokuiku/minna_navi/topics/topics4_02.html).
- [9] 村上亜由美、上島郁美、尾崎由美：食事時における白飯、おかずの食べ方と偏食との関連性、福井大学教育地域科学部紀要第 V 部 応用科学 (家政学編), Vol. 46, pp. 9–22 (2007).
- [10] 小林茂雄：鮮やかな光色で照明された食品に対する食欲、日本建築学会環境系論文集, Vol. 74, No. 637, pp. 271–276 (2009).
- [11] 小林茂雄、小口尚子：光色と BGM の種類がカフェでの会話行動に与える影響、日本建築学会環境系論文集, Vol. 71, No. 599, pp. 143–150 (2006).
- [12] 豊満美峰子、松本仲子：食物・食器・食卓の配色が嗜好に及ぼす影響、日本調理科学会誌, Vol. 38, No. 2, pp. 181–185 (2005).
- [13] 篠田孝祐、渡邊紀文、大山英明、大森隆司：ヒトにも走行性があるか: 光や音での行動誘導、龍谷大学.
- [14] 株式会社ロッテ：かむこと健康プロジェクト 咀嚼ランク表, <https://www.lotte.co.jp/kamukoto/partners/images/2022/10/34f74359842e453bd27b01d0b9c8a84e.jpg> (2022). 参照日 2025/1/24.
- [15] 近藤奈々美、小林茂雄：食品の種類と色彩から推測される食材と味覚、日本色彩学会誌, Vol. 43, No. 3, p. 141 (2019).
- [16] Hart, S. G. and Staveland, L. E.: Development of NASA-

- TLX (Task Load Index): Results of empirical and theoretical research, *Advances in psychology*, Vol. 52, Elsevier, pp. 139–183 (1988).
- [17] Brooke, J. et al.: SUS-A quick and dirty usability scale, *Usability evaluation in industry*, Vol. 189, No. 194, pp. 4–7 (1996).

# 脈拍変動を用いたデスクワークにおける 休憩方法による集中回復への効果検証

中島弥芳 <sup>†1</sup> 羽持涼花 <sup>†2</sup> ルペズ・ギヨーム <sup>†3</sup>  
青山学院大学 青山学院大学 青山学院大学

## 1. 序論

新型コロナウイルス感染症の蔓延をきっかけとしてオンライン授業やオンデマンド授業の導入が進んだ。Digital Knowledge 株式会社の調査で、オンデマンド授業を日常的に実施する大学は過半数を占めていると報告している [1]。また、全国大学生活協同組合連合が実施した調査によると、学生の約半数の大学生が学習に関する悩みを抱えており、授業や課題に対するモチベーション維持の難しさやリモートワークにおける集中力の低下が深刻な課題として挙げられている [2]。そのため、学習のための効果的な集中力回復法の検討が不可欠である。また長時間の集中作業は生産性の低下や肉体的および精神的疲労の原因となり得る。そのため、適切な休憩をとることにより、脳をリフレッシュさせ、集中力の持続や業務効率が向上することが期待される。本研究では、大学生を対象としたデスクワーク中の集中力回復に特化した最適な休憩方法を特定することを目的とする。具体的に、自律神経活動の客観評価である脈拍変動指標 LF/HF を用い、仮眠、音楽、ストレッチの 3 種類の休憩方法が集中力回復に与える効果を比較・検証することを目標とする。

## 2. 関連研究

### 2.1. ストレス・集中力評価と向上に関する研究

ウェアラブルデバイスを活用した心拍変動解析によるストレス評価や、集中力向上に関する手法は、これまで多く提案されてきている。

浦野らは、心理的緊張や興奮を可視化する手法について、ウェアラブルデバイスで計測した心拍数に基づいた光の変化が生じる LED デバイスの開発を行った [3]。これにより心拍波形を反映した発光パターンが心理状態の共有を可能とすることが確認され、心拍変動の医療分野やゲーム、イベ

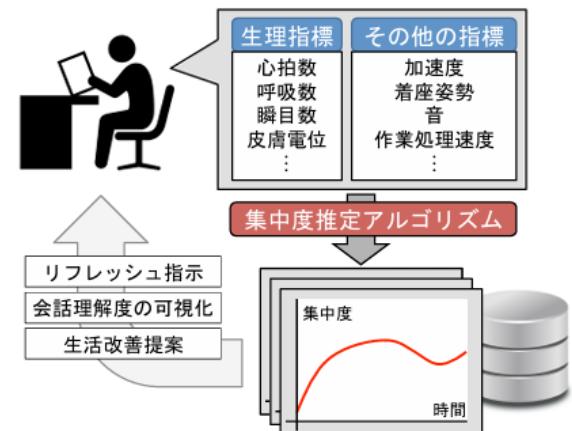


図 1 濱谷らの提案システムの構成図 ([5] より引用)

ントなどの応用可能性が示唆された。

井村らは、医療現場での医師のストレス推定手法について、腕時計型ウェアラブルデバイスを利用し、日中の心拍変動のデータや睡眠時における睡眠深度のデータから、日常生活の行動に対するストレスの状態を推定した [4]。そして、睡眠深度の変化や日中の心拍変動から算出した rMSSD および LF/HF のデータを基に疲労やストレス状態の確認を行った。その結果、医師のストレス状態を高精度で推定可能であり、ストレス軽減のための休息の必要性を判断できることが示された。

濱谷らは、スマートフォン内蔵センサやウェアラブルセンサを用いた集中度の推定手法を提案した [5]。図 1 は集中度推定結果に基づいてリフレッシュ等を促すシステムのアーキテクチャ図である。腕時計型ウェアラブルデバイスと衣服ポケットに収納したスマートフォンの内臓センサから取得されたタイミングや動画視聴などの行動中に取得した加速度、心拍数、皮膚導電率などのデータを分析し、集中度を推定した。1 名の被験者を対象に 40 分間のデータ収集を行い、集中行動とセンサデータの相関を分析した結果、集中時には特定の生理指標に特徴的な変化が見られることが確認された。

Conference Manuscript Format for Academy of Behavior Transformation by AIoT (BTI)

<sup>†1</sup> MIYOSHI NAKAJIMA, Aoyama Gakuin University

<sup>†2</sup> RYOKA HAMOCHI, Aoyama Gakuin University

<sup>†3</sup> GUILLAUME LOPEZ, Aoyama Gakuin University

## 2.2. 様々な休憩方法の効果に関する研究

渡部らは、大学生を対象に休憩時のクラシック音楽の聴取が休憩後の作業遂行能力や主観的気分に与える影響を調査した。符号課題をクラシック音楽聴取とノイズ音聴取の2条件で実施した結果、クラシック音楽条件での作業量が有意に多く集中力と活気が向上することが示された[6]。

荻野らは長時間作業の合間に実施する運動が集中力や作業効率に与える影響について調査した[7]。校正作業を休憩なし条件、運動なし条件、運動あり条件の3条件で実施し、生体情報と課題正答率の比較から検証を行った。結果、運動あり条件が正答率が最も高く、集中力維持に有効であることが示された。

三木らはオフィスワーカーにおける休憩方法と作業パフォーマンスに与える影響について、計算課題を課し、LF/HFと瞬きの回数から効果的な休憩手法の探索を行った[8]。休憩後の作業効率を評価した結果、「寝る」「運動をする」「タバコを吸う」が特にパフォーマンス向上に寄与することが確認された。しかし、課題の順番による慣れや個人差の影響がある可能性がある。また、休憩での5分間の睡眠は時間が短かったこと、および運動として採用したジョギングの負荷が大きすぎたことが挙げられた。

このように休憩時に仮眠や、音楽聴取、運動をすることでも休憩後のパフォーマンスが向上することが示されているが、従来研究の多くは実験環境や被験者属性が統一されておらず、デスクワーク時に特化した比較は十分ではない。

## 3. 本研究で用いる技術や用語について

### 3.1. 脈拍変動解析

自律神経系の活動には、心拍の呼吸と同期したゆらぎおよび、血圧と同期したゆらぎが反映されている。そのため、心拍間隔 (RRI) の周期的な変動に影響を与えていているという関係性が存在する。つまり、心拍変動 (HRV) を解析することによって、自律神経系の活動状態を把握することが可能であるといえる。また、交感神経系はストレスや、緊急事態において活性化され、副交感神経系は、脈拍数の減少や消化管の活動促進、エネルギー貯蔵の促進など、リラックス状態や回復を促進する。一方、心拍変動は心電図に現れるR波の間隔の変化から得られる指標であるため、日常生活での計測は難しい。その代替として、手首および、耳朶で計測可能な光電脈波から取得できる脈拍間隔 (PPI) の変化を表す脈拍変動 (PRV) を利用することが可能である。PRVは、HRVと少し違う値を示しているものの、同じく自律神経系の活動度合いを反映していることが示されている[9]。そのため、本研究では自律神経系の活動について脈拍変動解析を行い、休憩が与える影響の評価を行う。

時間領域解析は、脈拍間隔の変動を時間の観点から解析する手法であり、主に自律神経系の活動や脈拍の規則性を評価するために用いられる。特に、脈拍間隔が時間とともにどのように変化するかを明確に把握することに適しており、ストレス、疾患、運動負荷などの影響を解析する際に広く応用されている。周波数領域解析は、脈拍間隔の変動特性を周波数の観点から解析する方法であり、自律神経系の活動を詳細に評価するために広く用いられる。この解析手法では、脈拍間隔の変動を周波数として表現し、そのパワースペクトル密度を求ることで、交感神経系および副交感神経系の活動状態を定量的に評価できる。特に、低周波数帯域 (LF) と高周波数帯域 (HF) の比であるLF/HF比は、自律神経系の活動状態を定量的に評価する指標である。一般に、交感神経系の活動が副交感神経系の活動を上回る交感神経優位性を示すLF/HF比は、ストレス、疲労、精神的緊張の評価に活用される。本研究においてもストレス評価の指標としてLF/HF比を用いることとする。

### 3.2. 集中力の定義

先行研究より、LF/HF値が2以上の場合はストレス状態にあり、LF/HF値が2未満の場合はリラックス状態であると捉えることが一般的である[8]。集中力の観点においてもLF/HF値が2以上の場合は集中力が高い状態、2未満の場合は集中力が低い傾向がみられるとされている[10]。本研究では、デスクワークを行う際の集中作業時の状態を評価する。集中作業は精神的・認知的な負荷を伴うため、自律神経活動の観点から交感神経系が優位となり、ストレス状態として現れると解釈した。これにより、作業遂行に必要な精神的緊張を伴う状態 = 集中力が高い状態 (F/HF値が2以上) と定義し、この閾値を作業パフォーマンスの評価指標として検討した。

## 4. 休憩方法の効果検証実験

### 4.1. 比較対象の休憩方法

先行研究では、仮眠、運動 (ジョギング)、喫煙が作業パフォーマンスの向上に寄与することが示されている[8]。しかし、運動に関しては負荷が過大であると結論付けられていたため、本研究では軽いストレッチに置き換えることとした[11]。また、喫煙は非喫煙者に適用できない。その代替として、音楽を聴くという休憩方法を採用した。よって、本実験では、「仮眠」、「音楽を聴く（以下「音楽」）」、「軽いストレッチ（以下「ストレッチ」）」の3種類の休憩方法を比較した。休憩時間について、「仮眠」の場合のみ10分で、他の2つの休憩は3分とした。「音楽」と「ストレッチ」の休憩時間を3分としたのは、デスクワークの合間に現実的に



図 2 実験の様子 デスクワーク中（左上）/音楽を聴く（右上）/軽いストレッチ（左下）/仮眠（右下）

取り入れやすい時間であること、かつ安静時のベースライン測定時間と同一としたためである。また、「仮眠」の休憩時間に関しては、先行研究における 5 分という短時間設定では十分な効果が得られないと考え、先行研究を参考に休憩効果を高めるため 10 分へと延長した [12]。

#### 4.2. 実験手順

10 名の大学生（男 9 名、女 1 名）を対象に、静かな環境下において以下の手順で実験を実施した。被験者は実験説明を受け、実験に対する同意書による同意をもって、実験に参加した。図 2 は実験中の様子である。集中力への影響を考え、1 日 1 種類を 3 日間かけて行った。

- 椅子に座った状態でスマートウォッチを装着し、脈拍変動を計測する。
- 椅子に座って、3 分間安静状態を保つ。
- 20 分間の自由なデスクワークを行い、脈拍変動の記録を行う。
- 「音楽」(3 分)、「ストレッチ」(3 分)、「仮眠」(10 分)のいずれかの休憩方法をランダムで行う。
- 各休憩方法の後に再び 20 分間の自由なデスクワークを行い、脈拍変動を記録する。
- 3 回の実験後、実験アンケートの記入を実施する。

集中力回復効果は、安静時および、各作業中の脈拍変動から算出される LF/HF 比を用いて評価した。デスクワークについて、各被験者は前半の 20 分と後半の 20 分がそれぞれ同程度の負荷のものという条件のもと自由に行った。

#### 4.3. 脈拍数の解析手順

測定した脈拍数は以下の手順で解析した。まず、脈拍数から脈拍間隔を計算し、脈拍変動時系列データを 1 秒間隔にサンプリングした。フーリエ変換によりパワースペクトルを算出し、LF, HF, LF/HF をそれぞれ求めた。そして、各休憩方法について休憩前後における作業時の LF/HF 平均値の差を計算し、結果について t 検定を実施し、3 種類の休憩方法の効果と LF/HF 値に与える影響について、有意差の有無を判定した。

#### 4.4. 集中力回復の評価指標

##### 4.4.1 定量評価について

休憩後の LF/HF 平均値が休憩前の LF/HF 平均値と同レベルもしくは、上回ったときに集中力が回復・向上していると考えた。これは、休憩によって作業疲労や精神的緊張（ストレス）が一時的に軽減された後、再び集中作業に必要な高い精神的緊張状態へ速やかに移行できたことを示していると判断するためである。また、各休憩の組み合わせにおける LF/HF 値の有意差については、t 検定を行った。

##### 4.4.2 定性評価について

定性評価では 3 回の実験終了後実施した独自の実験アンケートを利用した（表 1）。アンケートでは、各休憩を実施後に集中力向上を感じたか、各休憩の休憩時間は適切だったか、各休憩中にリラックスすることができたか、また集中を妨げる要因として感じたもの、日常生活での取り入れやすさについて回答する質問を含む合計 10 個の質問で構成された。質問は「1. 全くそう思わない」から「5. とてもそう思う」の 5 段階での評価尺度を使用するもの（番号 1 から 9）、自由記述のもの（番号 10）を含む。

### 5. 休憩方法による効果比較の結果

#### 5.1. 脈拍変動指標による評価

3 種類の休憩方法の効果を比較するために、休憩前後の LF/HF の平均値を用いて、「仮眠・音楽」「仮眠・ストレッチ」「音楽・ストレッチ」の計 3 パターンの組み合わせについて、被験者内デザインであることを考慮し、対応のある t 検定を行った。

表 2 は各休憩方法間での t 検定の結果である。p 値についていずれの休憩方法間においても有意水準 ( $\alpha=0.05$ ) を超えたため、統計的有意差は認められなかった。また t 値と t 境界値についていずれの休憩方法間においても  $|t| < t$  境界値 の関係にあり、有意差は見られなかった。

表 3 は各休憩方法における、LF/HF 比の変化量と標準偏差である。休憩方法ごとの LF/HF の平均変化量は、「音楽」が 0.89、「仮眠」が 0.88、「ストレッチ」が 1.02 であり、

表 1 休憩方法に関する主観的評価アンケート項目

番号	質問内容
1	今回の実験において、休憩（音楽を聴く）後の方が休憩前よりも集中力が上がったと感じますか？
2	今回の実験において、休憩（ストレッチをする）後の方が休憩前よりも集中力が上がったと感じますか？
3	今回の実験において、休憩（仮眠をとる）後の方が休憩前よりも集中力が上がったと感じますか？
4	休憩（音楽を聴く）の設定時間は適切だったと感じますか。
5	休憩（ストレッチを聴く）の設定時間は適切だったと感じますか。
6	休憩（仮眠をとる）の設定時間は適切だったと感じますか。
7	休憩中（音楽を聴く）、リラックスできましたか
8	休憩中（ストレッチをする）、リラックスできましたか
9	休憩中（仮眠をとる）、リラックスできましたか
10	休憩方法に集中を妨げる要因（音や体制など）はありましたか。

表 2 各休憩方法間での t 検定の結果

項目	仮眠 vs 音楽	仮眠 vs ストレッチ	音楽 vs ストレッチ
p 値 (片側)	0.483	0.246	0.266
t 値	0.043	0.716	-0.650
t 境界値 (片側)	1.833	1.833	1.833
有意差	無し	無し	無し

表 3 休憩方法ごとの統計量

休憩方法	平均変化量	標準偏差
音楽	0.89	0.45
仮眠	0.88	0.47
ストレッチ	1.02	0.30

「ストレッチ」による覚醒が高くなる傾向が見受けられた。さらに、標準偏差に着目しても「仮眠」（標準偏差 = 0.47）、「音楽」（標準偏差 = 0.45）に比べて「ストレッチ」（標準偏差 = 0.30）が最も小さく、結果のばらつきが少なかった。

## 5.2. アンケートによる主観的評価

本実験では、独自のアンケートにより休憩方法に対する集中力向上の効果、設定時間の適切性、リラックス効果について主観的評価を行った。

### 5.2.1 各休憩方法の集中力向上効果についての評価

各休憩方法の集中力向上効果について、まず音楽を聴いた後に集中力が向上したと感じた被験者は、「とてもそう思う」が 20 %、「そう思う」が 70 % であった。一方で、「どちらでもない」と回答した被験者は 10 % を占めた。次にストレッチを行った後に集中力が向上したと感じた被験者は、「とてもそう思う」が 20 %、「そう思う」が 60 %、「どちらでもない」と「そう思わない」の回答がそれぞれ 10 % を占めた。最後に、仮眠をとった後に集中力が向上したと感じた被験者は、「とてもそう思う」が 40 %、「そう思う」が 60 % であった。よって各休憩方法の集中力向上効果について、「仮眠」は全被験者から「とてもそう思う」、「そう思う」という肯定的な回答を得られたと言える。「音楽」と「ストレッチ」に関しても肯定的な回答がそれぞれ全体の 90 %, 80 % を占めたといえる。

### 5.2.2 各休憩方法の設定時間についての評価

各休憩方法の設定時間について、まず音楽を聴く際の設定時間について適切だと感じた被験者は、「とてもそう思う」が 30 %、「そう思う」が 30 % であった。一方で「そう思わない」と回答した被験者が 40 % に上った。次にストレッチの設定時間について適切だと感じた被験者は、「とてもそう思う」が 40 %、「そう思う」が 40 % であった。一方で、「そう思わない」とする回答が 20 % を占めた。最後に仮眠の設定時間について適切だと感じた被験者は、「とてもそう思う」が 20 %、「そう思う」が 40 % であった。一方で「そう思わない」と回答した被験者も 40 % を占めた。よって、「音楽」と「仮眠」の設定時間について「適切だと思わない」という否定的な回答がどちらも 40 % に上り、「ストレッチ」についても否定的な回答が 20 % 見られたことから時間設定に課題があることが示唆された。

### 5.2.3 各休憩方法のリラックス効果についての評価

各休憩方法のリラックス効果について、まず音楽を聴いた際にリラックスできたと回答した被験者は「とてもそう思う」が 50 %、「そう思う」が 50 % であった。次にストレッチを行った際にリラックスできたと回答した被験者は「とてもそう思う」が 40 %、「そう思う」が 50 % であった。一方で「どちらでもない」とする回答が 10 % を占めた。最後に仮眠をとった際にリラックスできたと回答した被験者は、「とてもそう思う」が 50 %、「そう思う」が 40 % であった。一方で「どちらでもない」と回答した被験者が 10 % を占めた。よって「仮眠」は全被験者から「とてもそう思う」、「そう思う」という肯定的な回答を得られ、「音楽」と「ストレッチ」に関しても肯定的な回答が全体の 90 % を占めたと言える。

### 5.2.4 集中を妨げる要因についての評価

集中を妨げる要因について、各々自由記述で回答してもらった。その結果、休憩中に集中を妨げる要因としては仮眠時の

体勢や室内的明るさといった環境条件、スマートウォッチのバイブルーションや設定時間の短さが集中の妨げになっていることが課題として挙げられた。

## 6. 考察

休憩前後の LF/HF の平均値を用いて、t 検定を実施した結果、LF/HF 比の変化量には統計的に有意な差はなかった。一方で休憩方法ごとの LF/HF の平均変化量は、3 種類の休憩方法において「ストレッチ」がやや高い結果を示し、標準偏差に着目しても「ストレッチ」が最も小さく、結果のばらつきが少なかった。従って、「仮眠」、「音楽」、「ストレッチ」では、有意差が見られなかったものの、LF/HF の平均変化量が他の球威方法より高い傾向を示し標準偏差も最小であったことから、「ストレッチ」が比較的安定した効果を示す可能性が示唆される。有意差が見られなかったことについて、まず被験者 10 人と少ないうことが要因として考えられる。次に、短時間での休憩時間では LF/HF の十分な変化を検出できなかったことが要因として挙げられる。そのため測定方法の再検討や他の補助指標の併用を検討する必要があると思われる。

また、各休憩方法における集中力向上の評価では、「とてもそう思う」と「そう思う」の回答が多数を占めた一方で、「どちらでもない」または「そう思わない」とした回答が一定数存在した。リラックス効果に関する評価に関しても同様に、肯定的な回答が多数を占めたものの「どちらでもない」とした回答も見られた。これより、各休憩方法の効果には被験者間での個人差が大きく、統一的な効果を示すことが難しかったのではないかと考えられる。

更に、仮眠の体勢や室内的明るさといった物理的要因や休憩の設定時間といった実験設計上の問題も、休憩効果を低下させたと考えられる。以上を踏まえ、サンプルサイズを 20 人以上に拡大することや、個人差を考慮した評価指標や実験方法の検討、さらに実験設定の見直しを検討する必要がある。また今回は t 検定のみ行ったが、反復測定分散分析といった別の統計手法による解析も検討する必要があると考えられる。

## 7. 結論と今後の展望

本研究では、大学生を対象にデスクワーク中の集中力回復に特化した最適な休憩方法を特定することを目的として、仮眠、音楽、軽いストレッチの 3 種類の休憩方法が集中力回復に与える効果を比較・検証した。実験では 20 分のデスクワークの後に 3 種類からランダムに休憩を行った。そして再度デスクワークを行った後に安静時および脈拍変動から LF/HF を算出し、集中力回復効果を比較・評価した。被験

者の LF/HF スコアから t 検定を行った結果、統計的に有意な差があるとは言えなかった。

一方で、休憩方法ごとの LF/HF の平均変化量からストレッチによる覚醒が高くなる傾向が見受けられること、結果のばらつきが少ないことが分かった。休憩方法の集中力向上・リラックス効果や、設定時間の適切性に関するアンケートの結果から、仮眠が全被験者から肯定的な回答を得られた。

今後は、パネラー数を増やすことや、休憩方法ごとの効果を最大化するために事前調査やパイロット実験を行い、最適な時間設定を再検討する必要がある。また、仮眠時の体勢や光量、音楽の再生時間や音量の調整などといった環境要因の見直しも必要である。休憩効果には個人差が大きいことが示唆されているため、クラスター分析や多変量解析を用いたデータ分析を通じて被験者の特性ごとに適した休憩方法を明らかにしたい。さらには主観的評価に加えて他の生理指標を併用した多角的な効果検証を行うとともに、反復測定分散分析による解析も行い、統計的妥当性を高めていきたい。

## 参考文献

- [1] DigitalKnowledge 株式会社：オンデマンド学習に関するレポート、[https://www.digital-knowledge.co.jp/wp-content/uploads/2023/11/report\\_ondemand.pdf](https://www.digital-knowledge.co.jp/wp-content/uploads/2023/11/report_ondemand.pdf). (参照日 2025/01/28).
- [2] 全国大学生活協同組合連合会：第 59 回学生生活実態調査報告、<https://www.univcoop.or.jp/press/life/report.html>. (参照日 2025/01/28).
- [3] 浦野健太, 廣井慧, 米澤拓郎, 河口信夫ほか：ドキドキをセンシングして可視化する LED ライティングデバイス, マルチメディア, 分散協調とモバイルシンポジウム 2248 論文集, Vol. 2020, pp. 1616–1622 (2020).
- [4] 伊村一成, 後藤佑介, 酒井晃二, 小原雄, 田添潤, 三浦寛司, 廣田達哉, 内山彰, 乃村能成ほか：ウェアラブルセンサデバイスを用いた医師のストレス推定手法の提案, 研究報告電子化知的財産・社会基盤 (EIP), Vol. 2021, No. 5, pp. 1–8 (2021).
- [5] 濱谷尚志, 内山彰, 東野輝夫ほか：種々のセンサを併用した集中度センシング法の検討, 研究報告高度交通システムとスマートコミュニティ (ITS), Vol. 2015, No. 10, pp. 1–6 (2015).
- [6] 萩野隼：休憩時の運動が生体情報と集中力に及ぼす影響, 日本知能情報ファジィ学会第 36 回ファジィシステムシンポジウム講演論文集, pp. 35–36 (2020).
- [7] 渡部貴史, 小林未来：休憩時の音楽聴取が作業遂行に及ぼす影響, 新潟医療福祉学会誌, Vol. 20, No. 1, pp. 39–39 (2020).
- [8] 三木隆裕, 他：休憩時間の過ごし方が作業パフォーマンスに及ぼす影響の調査, 研究報告ユビキタスコンピューティングシステム (UBI), Vol. 2018, No. 3, pp. 1–8 (2018).
- [9] Yuda, E., Shibata, M., Ogata, Y., Ueda, N., Yamabe, T.,

- Yoshizawa, M. and Hayano, J.: Pulse rate variability: a new biomarker, not a surrogate for heart rate variability, *J Physiol Anthropol*, Vol. 39, No. 21 (2020).
- [10] 中川千鶴：特集3：人間工学のための計測手法 第4部：生体電気現象その他の計測と解析 (5)-自律神経系指標の計測と解析, 人間工学, Vol. 52, No. 1, pp. 6–12 (2016).
- [11] Work, H.: 集中力を高める休憩の取り方：生産性を上げるためのポイント, <https://note.com/hiromoneywork/n/n1202c250d8e0>. (参照日 2025/01/23).
- [12] G-NAP: 仮眠の効果 ~15~20 分程度の仮眠で疲労回復効果~, <https://g-nap.com/post-987/>. (参照日 2025/01/23).

# 行動変容における両価性を乗り越えるための動機づけメッセージ生成手法に関する選好調査

有賀玲子<sup>†1</sup> 片桐有理佳<sup>†2</sup> 佐藤妙<sup>†3</sup> 藤村香央里<sup>†4</sup> 渡邊淳司<sup>†5</sup>  
NTT 株式会社

## 1. はじめに

近年、ユーザ情報に基づいてパーソナライズされた、その人にとって望ましい行動を促す情報人々が接する機会が増えている。しかし人はしばしば、「行動を変えたほうがいい」という理性的な認識と、「行動を変えることは気が進まない」という直感的な感情を同時に併せ持つ両価的な心理状態に陥り、行動の価値を認識したとしても結局行動には移さないことがある。意図・行動ギャップに関する研究[1]では、行動する意図を持ったとしても、わずか半数しか実際の行動に移さないことが確認されている。特に、将来の疾病発症を予防するための健康行動（例えば、身体活動の増加や食習慣の改善など）は、その効果が直ちに現れないため、先延ばしにされがちである[2]。しかし、健康行動の先延ばしは、疾病の発症・重症化などの重大な結果を招く恐れがある。生活習慣が発症の大きな要因である糖尿病や高血圧などの疾患の有病率は、過去 20 年間で世界的に増加しており[3]、疾病予防のための動機づけ支援は世界的な社会課題となっている。このため、取り組んだほうがいいと理解していても気が進まない活動（Reluctant Activities: RA）を動機づける新たな手法を検討することは重要である。

多くの研究で、健康行動を動機づける方法が検討されてきた。最新の研究では、明確な目標を持ち両価的ではない個人については、生成 AI を用いた対話型エージェントが動機づけに有効であることが確認されている[4]。しかし、両価的な層については、生成 AI による動機づけ支援の実践的な指針がほとんど報告されていない。両価的な人の心理は複雑で、彼らを動機づけることは容易ではなく、丁寧な設計が必要である[5]。そこで本研究では、両価的な人に特化した生成 AI による動機づけ支援の実践的な調整の指針を得るために、行動科学とヒューマンコンピュータインタラクション (HCI) の文献レビューに基づき、両価的な人の複雑な心理を整理し、2 つのモチベーションプロファイルを導出した。そして、各プロファイルに含まれる行動の抑制要因に対処する 2 種類の動機づけメッセージの設計を行った：ひとつは、RA 自体の魅力を高め、肯定的な感情（楽しそう、

好奇心が湧く等）を増幅させる方法（ポジティブ感情促進型）、もうひとつは、RA に対する否定的な感情（大変そう、やる時間がない等）を軽減させる方法（ネガティブ感情軽減型）である。

本稿では、両価的な人におけるこれらのメッセージ方式に対する選好調査の結果を報告する。本調査は、健康行動に取り組んだほうがいいと理解しているが気が進まないと回答した成人を対象とし、オンラインで実施された。調査では、気が進まない活動と好きな活動を取得し、これを生成 AI に入力することで参加者ごとに個別に 2 方式の動機づけメッセージを 12 個ずつ生成した。参加者は、合計 24 個のメッセージの中から最も動機づけられたメッセージを選択し、これについて内発的動機づけに関する印象評価を行った。2~3 週間後の追跡調査では、選択したメッセージに書かれていた行動を実施したかどうか調査を行った。有効回答数 548 件の調査結果から、気が進まない活動に対する「自己効力感」と「気軽さ」の認識が低い人ほど、ポジティブ感情促進型メッセージを好む傾向にあることが分かった。また、追跡調査の結果から、どちらのタイプを選択した人においても、メッセージの内容を実行した人の割合は、行動変容ステージにおける無関心期において約 20%，関心期において約 50%，準備期においては 60% 以上であった。これらのこととは、気が進まない活動に対する認識の違い（すなわち、両価的な態度の原因となっている抑制要因の違い）がメッセージ選好に影響を与えること、また、提案する動機づけメッセージの設計が両価的な人を実際の行動に至るまで動機づける可能性を示唆している。

## 2. 関連研究

### 2.1 健康に関する行動変容の理論と介入

健康に関する初期の行動変容研究では、公衆衛生の分野において、人々が疾病予防プログラムになぜ参加しないのかを説明する心理のモデル化が注目を集め、健康信念モデル[6]、社会的認知理論[7]、行動変容ステージモデル[8]などの行動変容を支援するための理論的枠組みが活発に開発された。また、これらのモデルに基づき、行動変容に対して両価的な人を動機づけるためのカウンセリング技法である動機づけ面接[9]も確立され、アルコール依存症、禁煙プログラム、体重管理など、現在に至るまで幅広く健康分野で活用してきた。

その後、テクノロジーの高度化に伴い、行動変容の研究

Preference Survey on Motivational Message Generation Techniques to Overcome Ambivalence in Behavioral Change

<sup>†1</sup> REIKO ARUGA, NTT, Inc.

<sup>†2</sup> YURIKA KATAGIRI, NTT, Inc.

<sup>†3</sup> TAE SATO, NTT, Inc.

<sup>†4</sup> KAORI FUJIMURA, NTT, Inc.

<sup>†5</sup> JUNJI WATANABE, NTT, Inc.

は、「なぜ人は行動を起こさないか」を説明する理論的枠組みから、「どのようにすれば人は行動を起こすか」という介入に関する研究が盛んに行われるようになった。モバイル端末やウェアラブルデバイスの登場により、介入の内容やタイミングが細かく設計可能となり、身体活動、食習慣改善、睡眠改善など、様々な健康分野での介入の有効性が確認してきた。さらに Michie ら[10]は、多様な介入手法を分析し、介入の最小構成要素を特定した。これにより、介入手法の構成要素の記述が標準化され、介入を詳細に設計・評価することが可能となった。

しかし、理論や介入に関する研究が充実してきたにも関わらず、両価的な人を動機づけることは未だ困難であり、彼らを動機づけることに特化した心理モデルや介入の指針は確立されていない。そこで本研究は、両価的な人を動機づけるための介入設計の初期指針を得ることを目指す。具体的には、従来の理論と介入に関する知見を文献調査に基づき統合することにより、両価的な人が抱える行動の抑制要因を細分化し、各要因に対処するための介入手法の設計と評価を行う。

## 2.2 両価的な人の動機づけ

両価的な人を動機づける手法のひとつに、動機づけ面接がある。動機づけ面接は、行動を変えたい・変えたくないという両価的な人の抱える複雑な心理に寄り添い、行動を変容させたいという気持ちを高める手法で、禁煙、アルコール依存症、運動不足など、幅広い分野で活用してきた。近年では、LLM の高度化により、動機づけ面接やカウンセリングのチャットボット化がさかんに検討されている。一方で、先行研究により報告されている動機づけ面接による改善率にはばらつきがあり、その効果は未だ確実なものとは言えない。

行動変容ステージごとに LLM による動機づけの効果を検証した研究では、実行期以降の目標が明確な層（すなわち両価性が低い層）には有効だが、両価性を示す層には効果が確認されず、この層の心理の複雑さが要因であると指摘されている[4]。さらに、Lee ら[11]は従来の HCI 専門家による行動変容システムの設計では動機づけの十分なパーソナライズが困難であるとし、ユーザ主体の計画立案支援の有効性を報告している。これらを踏まえ、我々は両価的な人のモチベーションプロファイルをより詳細に捉え、これに基づき動機づけをより一層パーソナライズするアプローチが必要であると考えた。

## 2.3 行動の生起要因に関するモデル

両価的な人が抱える行動の抑制要因を細分化し、モチベーションプロファイルの仮説を立てるにあたり、本研究では COM-B モデルを参照する。COM-B モデル[12]は、人間の行動がどのような条件で生起するかを統合的に説明し、介入設計に応用可能である実用的なフレームワークとして、公衆衛生の分野で広く活用されており、患者への介入やパ

ンデミック時の政策の設計などに応用されている。COM-B モデルでは、行動 (Behaviour) の生起要因として、3 つの要素（以下、行動の構成 3 要素）を定義している：能力 (Capability)、機会 (Opportunity)、動機 (Motivation) である。動機はさらに細分化され、計画、評価、信念などの意識的・理性的な思考を伴う「内省的動機」と、連合学習、模倣学習、生得的気質などから形成される無意識的・直感的な「自動的動機」が定義されている。この分類は、二系統の認知処理に関する心理学の知見と合致している[13]。

COM-B では、モデルの定義だけでなく、様々な介入との接点が網羅的に示されているが、具体的な行動の生起に関するシステムデザインの要件は示されていない。次節では、両価性の高い人に対する介入を設計する上での留意点について、体験デザインの観点から整理する。

## 2.4 行動の生起率を高める介入手法

動機づけや習慣形成を支援するために、説得技術に関する HCI 研究では数多くのアプローチが模索されてきた。その中でも、Fogg の Behaviour Model (FBM) [14]は、ユーザの具体的な行動を誘発する体験デザインの場面で広く有効性が実証してきた。FBM は、モチベーションの低い層に対しては、タスクを簡素化し、知覚される能力を高めることが効果的であると強調し、簡素化の 6 要素を定義している：時間、コスト、身体的努力、認知的努力、社会的規範、ルーティン[14]である。

一方で、内発的動機づけに関する研究では、行動の持続における「楽しさ」の役割が重要視されてきた[15]。Milkman らは、すべき活動（例：ジム通い）と楽しい活動（例：人気オーディオブックの視聴）を組み合わせることで継続率が向上することを実証し、誘惑バンドリングの有効性を示した[16][17]。誘惑バンドリングをシステムで自動的にパーソナライズすることは困難であったが、本研究では、LLM を活用した誘惑バンドリングのパーソナライズの実現可能性を探る。

目標行動の簡素化と誘惑バンドリングは、それぞれ介入の手法として有効性が示されているにも関わらず、両価性の高い人の動機づけメッセージの設計として十分統合されていない。本研究では、COM-B における行動の構成 3 要素の観点から、両価的な人が抱える RA に対する行動の抑制要因を整理し、モチベーションプロファイルの仮説を立て、各要因に対処する介入を設計することを試みる。

## 3. 動機づけメッセージの設計

我々は、COM-B モデルに基づき、「取り組んだほうがいい」と理解しているが気が進まず行動に移せない」という両価的な心理を整理し、2 つのモチベーションプロファイルを導出する。続いて、前章で述べた行動科学や HCI の文献レビューに基づき、これらのプロファイルに含まれる行動の障壁に対処する 2 種類の動機付けメッセージを設計する。

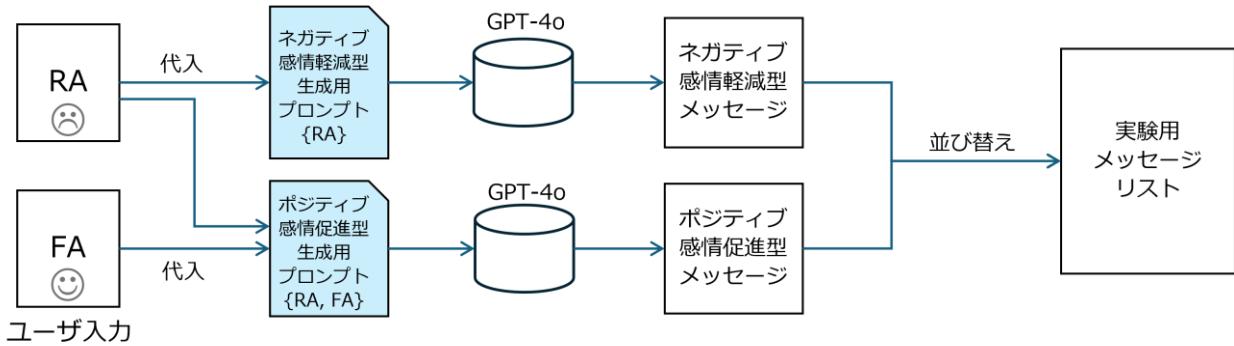


図1 実験用メッセージ生成システム

表1 生成されるメッセージの例

メッセージタイプ	RA	FA	メッセージ例
ネガティブ感情軽減型	ウォーキング	-	最初から長時間歩く必要はありません。まずは1日5分だけ、家の周りを軽く歩いてみましょう。
	ランニング	-	毎日でなくても大丈夫。まずは1週間に1回だけ走ることを目標にしてみましょう。無理のないペースで日常に取り入れることができます。
	筋トレ	-	その日の体調や気分に応じて、筋トレの回数を調整しましょう。無理せず少ない回数でも、自分に合ったペースで続けることが大切です。
ポジティブ感情促進型	ウォーキング	推し活	推し活を楽しみながら、散歩中にお気に入りの推しの曲を聴くことで、気分も良く、足取りも軽くなります。
	ランニング	犬と遊ぶ	愛犬とキャッチボールを楽しむ際、ボールを取りに行く過程を軽いジョギングに変えることで、自然にランニングを取り入れることができます。
	筋トレ	ガーデニング	ガーデニングをしながら、土を運ぶ際には小さな鉢を使って繰り返し持ち運びを行うと、腕や背中、肩の筋肉を鍛えることができます。

### 3.1 両価的な人の適切な動機づけ

以下では、RAを抱えるユーザの心理状態をCOM-Bモデルの3要素から整理する。彼らはRAについて「取り組むべき」と認識しており、内省的動機はある程度存在するものと考えられる。一方で「気が進まず行動に移せない」のは、自動的動機が抑制的に働き、無意識的・直観的に行動を回避しているためと考えられる。このような両価性を示す個人には、単なる説得は有効ではないことが知られている。次節では、直感的抑制の要因に基づき、2つのモチベーションプロファイルの仮説を提示する。

### 3.2 ネガティブ感情軽減メッセージ

1つ目のモチベーションプロファイルは、能力または機会の不足による抑制に関するものである。COM-Bモデルによれば、行動の構成3要素は相互に作用するものであり、能力または機会の不足は、動機を低下させる。つまり、ユーザがRAについて「要求が厳しすぎる」または「機会が不足している」と認識している場合、自動的動機が抑制的に働く。こうした場合、これらの認知的障壁を低減し、ネガティブな感情を軽減することで動機が高まる可能性がある。

FBMで説明されているタスクの簡素化は、この課題に直接

対処するものである。我々はFBMに基づき、RAをより小さなタスクに再構成したメッセージをLLMにより生成し、これをネガティブ感情軽減型メッセージとした。

### 3.3 ポジティブ感情促進型メッセージ

2つ目のモチベーションプロファイルは、RAがネガティブな感情と結びついている場合である。これはオペラント条件付け[18]で説明できる。オペラント条件付けは、事象と結果の間の関係性を学習するプロセスであり、好ましい結果が得られる行動は起こりやすく、好ましくない結果が得られる行動は回避されやすくなることが示されている。例えば、ランニングが「苦しいだけ」と学習されれば、理性的な説得だけでは行動に至りにくい。このような場合、ネガティブな感情を好ましい感情に置き換えることが有効であり、誘惑バンドリングはその直接的な方法である。本研究では、このプロファイルに対処するため、ユーザの好きな活動(Favorite Activity: FA)の一部をRAで置き換える形でパーソナライズした誘惑バンドリングをLLMによって生成し、これをポジティブ感情促進型メッセージとした。各メッセージの詳細な生成過程は次章で述べる。

## 4. 動機づけメッセージ生成

我々は、ポジティブ感情促進型メッセージおよびネガティブ感情軽減型メッセージから構成される実験用メッセージのリストを GPT-4o を用いて生成した。メッセージ生成の前処理として、ユーザから RA および FA を取得する。本章では、RA 及び FA を取得するための質問紙設計と、メッセージ生成手順について述べる。

### 4.1 RA 取得質問紙

本研究では健康のための身体活動を動機づけることに焦点を当てる。WHO は、18 歳から 64 歳の成人に対して、中強度の有酸素運動、強度の高い有酸素運動、および筋力強化運動を健康のために行うことを推奨している[19]。これを踏まえ、それぞれの代表的な活動としてウォーキング、ランニング、筋力トレーニングを RA の選択肢として選定した。RA 取得質問紙は、以下の通りである：

取り組んだ方がいいと思うけれど気が進まない運動習慣を次の中から 1 つ選んでください。

- (1) ウォーキング (2) ランニング (3) 筋力トレーニング

### 4.2 FA 取得質問紙

ポジティブ感情促進型メッセージは、個人の FA を RA と組み合わせて生成する。我々は、より強い動機づけを促すため、モチベーション及び快の情動と関連する生体反応に着目した。神経科学の文献によれば、自発的な行動や報酬による快感時には共通の神経系が駆動し、ドーパミンが分泌されることが知られている [20]。そこで、ヒトでドーパミン分泌が確認された実験課題を参照し、喚起される感情を整理した。例として、ビデオゲーム [21]、金銭的報酬の予測 [22]、恋人の写真提示 [23] などがあり、これらは楽しさ、興奮、期待感など活性度の高いポジティブ感情を誘発するものと考えられた。我々は、このような感情を喚起する FA を「人が自発的に取り組み、活性度の高いポジティブ感情を感じる活動」と定義した。生成されるメッセージが日常的に実行可能である必要があることを考慮し、以下の 2 項目を FA 取得質問紙とした：

(1) 普段の生活の中でできる、好きなことや楽しいことを教えてください。

(2) 仕事がある日でもできる、気分転換やリフレッシュになることを教えてください。

ここで、FA 取得質問紙は、ユーザ固有の体験を具体的に捉るために自由記述形式とした。

### 4.3 メッセージ生成過程

メッセージ生成の処理フローを図 1 に示す。メッセージの生成には GPT-4o [24] を用いた。メッセージ生成のためのプロンプトは、実験者の意図通りのメッセージが出力されるように試行錯誤を重ねて設計された。最終的なプロンプトは、ポジティブ感情促進型では、「ユーザの好きな活動 (FA) と同時に RA を行う方法を提案せよ」、ネガティブ感

情軽減型では、「RA を行うためのスマルステップを提案せよ」というものであった。取得した RA と FA は、メッセージ生成用のプログラム中でプロンプトに代入された。生成されたメッセージの例を表 1 に示す。

ポジティブ感情促進型メッセージについては、入力となる好きな活動 FA は人によって異なるため、参加者ごとに個別に生成するものとした。一方、ネガティブ感情軽減型メッセージについては、同じ RA を選択した参加者については同一のメッセージ群を評価に使用した。なぜなら、FBM が示す「簡素化の 6 要素」は網羅的であり、これに基づくメッセージは多くの人にとって共通して「大変さ」を軽減する内容であると考えられるためである。

ネガティブ感情軽減型メッセージ群の生成手順として、まず、ウォーキング・ランニング・筋トレの各 RA について、ネガティブ感情緩軽減メッセージ生成プロンプトにより複数のメッセージを生成した。続いて、生成されたメッセージの中から、簡素化の 6 要素について考慮されているものを、要素が偏らないようにバランスよく抽出した。具体的には、時間的拘束の短い活動（“家の周りを 5 分走る”等）、知覚される身体的負荷が低い活動（“まずは一種類の筋トレからはじめよう”等）、認知的負荷の低い活動（“毎日できなくてもよい”等）をバランスよく選択した。また、選択したすべてのメッセージの内容について、社会的規範および日常性を逸脱していないか、すなわち、社会のルールから外れるものでもなく、日常から大きく外れるものでもないことを確認した。こうして 12 個のメッセージを抽出し、これを否定的感情軽減型メッセージセットとした。

最終的に、ポジティブ感情促進型 12 個（参加者ごとに生成）とネガティブ感情軽減型 12 個（RA 別に共通）をランダムに並び替えた 24 個を実験用メッセージリストとした。

## 5. 調査手順

2 種類の動機づけメッセージデザインに対する選好パターンを調査するために、オンラインアンケートを実施した。我々の一連の研究は、事前調査、本調査、追跡調査で構成された。事前調査では、RA と FA についてアンケートを行った。続いて、事前調査で得られた RA と FA を、4 章で示したメッセージ生成システムに入力し、参加者一人一人に実験用メッセージリストを作成した。本調査では、生成されたメッセージのうち最も動機づけられたメッセージを 1 つ選択し、印象評価を行った。本調査から 2~3 週間後に実施された追跡調査では、本調査で選択したメッセージの内容を実行したかどうかアンケートを行った。

### 5.1 参加者

調査は、100 万人程度のモニタを抱える日本の調査会社を通じて実施された。本研究は研究所の倫理委員会による審査・承認を得た。参加者は 20~50 歳の就労成人であり、参加は任意であった。調査対象は、ある活動を行うべきだと

表 2 内発的動機づけに関する評価項目

ID	COM-B	IMI	評価項目
C-1	Capability	知覚される能力	私はこれができる
C-2		知覚される能力	私の能力やスキルに合っている
C-3		努力	気軽な取り組みだ
O	Opportunity	努力	普段の生活に無理なく取り入れることができる
M-1	Motivation (automatic)	興味・楽しみ	わくわくする
M-2		興味・楽しみ	楽しそうだ・興味深い
M-3	Motivation (reflective)	価値・有用性	これをすることは多少なりとも私にとってプラスになる
M-4	Motivation (total)	-	やってみたい

表 3 参加者の属性

N = 548	
年齢	20-24 (3%), 25-29 (16%), 30-34 (11%), 35-39 (16%), 40-44 (11%), 45-49 (16%), 50-54 (15%), 55-59 (11%)
性別	男性 (52%), 女性 (48%)
行動変容ステージ	無関心期 (27%), 関心期 (39%), 準備期 (34%)
選択した RA	ウォーキング (29%), ランニング (29%), 筋トレ (42%)

理解しているが、先延ばしがちな人（両価的な人）とした。調査前にスクリーニングを実施した。適格基準は以下の通り：(i)重要と認識しながら先延ばしがちな活動 RA（ウォーキング、ランニング、筋力トレーニング）を少なくとも 1 つ有すること、(ii)選択した活動について行動変容ステージモデルのうち「無関心期」「関心期」「準備期」のいずれかの段階にあること（実行期・維持期を除く）。調査会社のシステムは、年齢・性別・RA のバランスを取りつつ、これらの基準に基づき参加者を自動的に選定した。調査会社は完了した調査数に応じて謝礼を支払った。

## 5.2 データ収集

我々は 2024 年 9 月から 10 月にかけて一連の調査を実施した。参加者は調査会社のシステムを通じて募集され、同意後にスクリーニング条件に基づき自動的に選定された。調査対象者として選定された者には、事前調査が配布された。回答期限は 7 日間とした。事前調査から 14 日後に本調査を、さらに 14 日後に追跡調査を実施し、期限はそれぞれ 4 日間と 6 日間であった。

調査は任意参加であり、回を追うごとに離脱が生じる可能性があった。追跡調査の有効回答数を確保するため、事前調査への参加者として、追跡調査の最低目標の約 5 倍となる 1319 名を通過させた。結果として、本調査で 647 件、追跡調査で 499 件の回答を得た。スクリーニング調査では、年齢・性別・年収区分などの人口統計情報に加え、RA に関する行動変容ステージを収集した。事前調査では、4.1 節で示した 2 種類の FA 取得質問紙に自由記述形式で回答した。さらに、RA としてウォーキング、ランニング、筋力トレーニングのいずれか 1 つを選び、選択した RA に対する印象を 7 件法リッカート尺度（「6：とてもそう思う」から「0：まったくそう思わない」まで）で評価した。ここで、RA に対する印象評価の項目は、確立された内発的動機づけ尺度（Intrinsic Motivation Scale: IMI）[25] [26] [27] [28] を参考に、COM-B モデルの 3 つの行動要素に対応付けて作成した（表 2）。

ニングのいずれか 1 つを選び、選択した RA に対する印象を 7 件法リッカート尺度（「6：とてもそう思う」から「0：まったくそう思わない」まで）で評価した。ここで、RA に対する印象評価の項目は、確立された内発的動機づけ尺度（Intrinsic Motivation Scale: IMI）[25] [26] [27] [28] を参考に、COM-B モデルの 3 つの行動要素に対応付けて作成した（表 2）。

事前調査の回収後、4 章で示したメッセージ生成の手順に従い、24 個のメッセージから成る実験用メッセージリストを参加者ごとに作成した。本調査で参加者は、実験用メッセージリストの中から最も動機づけられると感じたメッセージを 1 つ選択した。その後、選択したメッセージに対する印象評価を行い、メッセージの良かった点について自由記述で回答した。印象評価の内容は、RA に対するものと同一とした。追跡調査では、本調査で選択されたメッセージが再度アンケート画面に提示され、参加者は本調査後にその内容を実際に実行したかどうかを報告した。

## 5.3 データ分析

内発的動機づけの 8 項目について、当初気が進まない活動として選択された RA と、参加者が最も動機づけられると選択したメッセージを比較するため、メッセージタイプごとに対応のある t 検定を行った。各項目について、まず ANOVA と等分散性検定を行い、結果に応じて等分散または不等分散の t 検定を適用した。多重比較にはボンフェローニ補正を用いた。さらに、2 種類のメッセージタイプの心理的効果を比較するため、RA とメッセージ評価の差分について同様の手順で分析した。

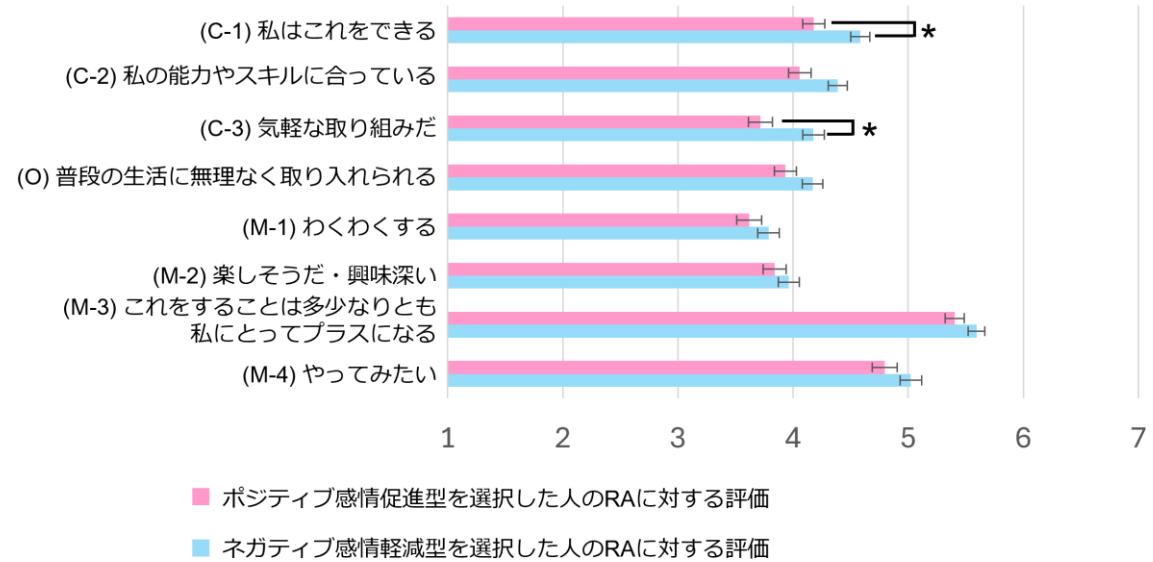


図 2 RA に対する内発的動機づけ評価の平均

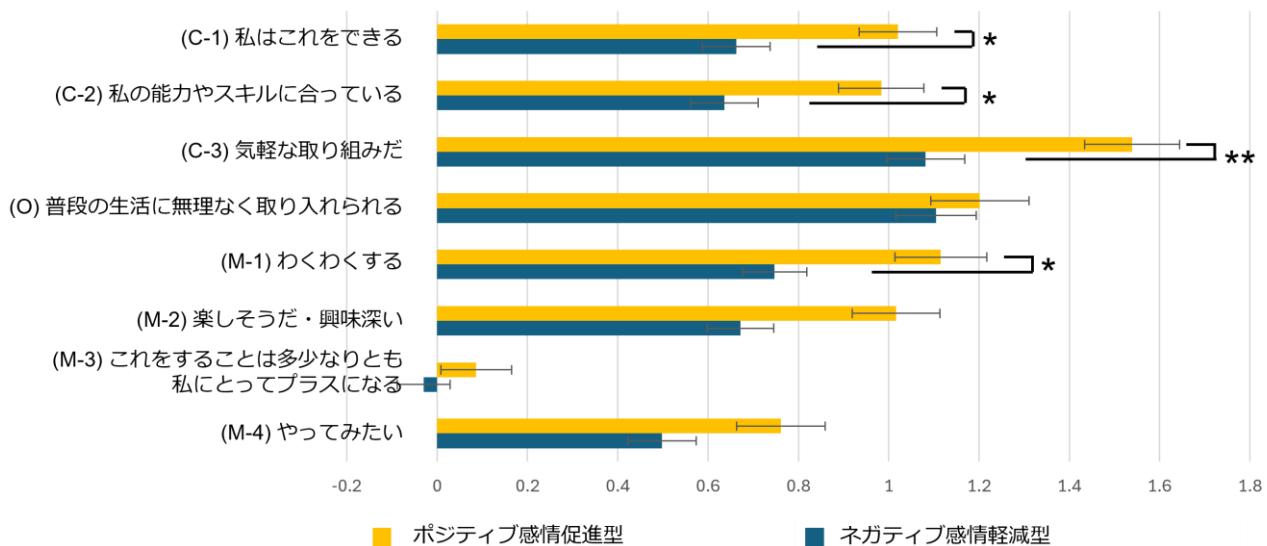


図 3 RA と選択されたメッセージに対する内発的動機づけスコアの差分

また、メッセージ選好と性別、行動変容ステージ、RA の種別との関連性を確認するため、各属性とメッセージ選好の間に有意な関連があるかをカイ二乗検定で評価した。有意水準は  $p < .05$  とした。

## 6. 結果

分析の前に、回答の信頼性を担保するために、本調査と追跡調査において回答がストレートライニングであった者を除外した。このフィルタリングにより、本調査については 548 件、追跡調査については 407 件を分析対象として抽出した。参加者の年齢、性別、RA、行動変容ステージは表 3 に示す通り多様であった。

### 6.1 メッセージ選好とユーザ属性の関係

本調査の結果、548 名のうち、305 名がネガティブ感情軽減タイプ、243 名がポジティブ感情軽減タイプを、最も動機づけられるメッセージとして選択した。RA に対する内発的動機づけの 8 項目を比較したところ、ポジティブ感情促進型メッセージを選択した参加者は、ネガティブ感情軽減型を選択した参加者よりも「私はこれをできる」「気軽な取り組みだ」の 2 項目で有意に低い得点を示した ( $p < .05$ ) (図 2)。また、カイ二乗検定の結果、メッセージ選好は、性別 ( $\chi^2 (1) = 0.52, p = 0.77$ )、行動変容ステージ ( $\chi^2 (2) = 3.79, p = 0.15$ )、気が進まない活動種別 ( $\chi^2 (2) = 2.46, p = 0.29$ ) のいずれにおいても統計的な関連性は認められなかった。

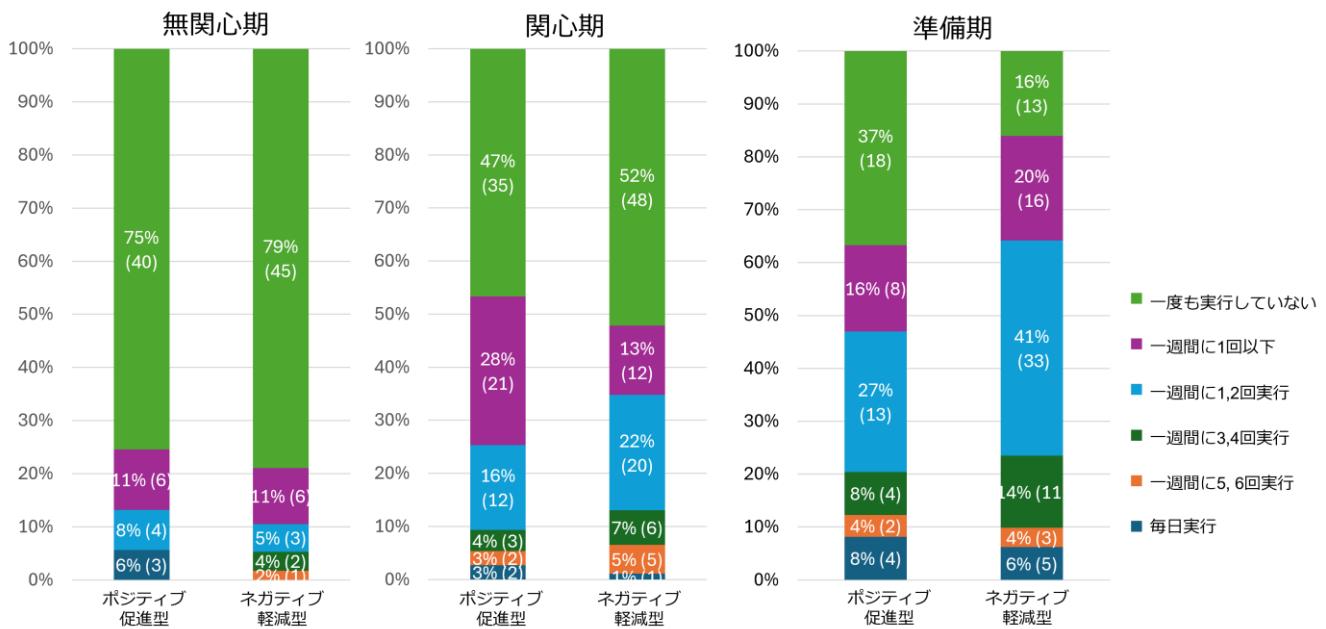


図 4 追跡調査の結果

## 6.2 メッセージデザインによる動機づけ効果

RA と選択したメッセージの内発的動機づけ 8 項目を選択したメッセージタイプ間で比較したところ、どちらのタイプにおいても「これをすることは多少なりとも私にとってプラスになる」を除く全項目で、RA よりも選択されたメッセージの方が有意に高いスコアを示した ( $p < .05$ )。

選択されたメッセージにより、RA に対する内発的動機づけに関連するどの認識がどのように変化したか確認するため、内発的動機づけの 8 項目について、RA と選択されたメッセージの差分を算出し、メッセージタイプ間で比較した(図 3)。その結果、以下の 4 項目でポジティブ感情促進型の方がネガティブ感情軽減型よりも有意に高かった: 「気軽な取り組みだ」( $p < .01$ )、 「私の能力やスキルに合っている」( $p < .05$ )、 「私はこれをできる」( $p < .05$ )、 「わくわくする」( $p < .05$ )。

追跡調査の集計結果を図 4 に示す。行動変容ステージが上がるごとに、一度でも実行した人の割合が増えていることが確認された。

## 7. 考察

メッセージ選好は性別や行動変容ステージとは関連せず、RA に対する内発的動機づけと関係していた。知覚される能力と気軽さの認識が低い参加者はポジティブ感情促進型を動機づけメッセージとして好む傾向が見られた(図 2)。また、ポジティブ感情促進型に動機づけられた群は、ネガティブ感情軽減型に動機づけられた群と比較して、もともとの RA に対する気が進まない状態の心理から、COM-B における「能力」に関する項目と「自動的動機」に関する項目が有意に高められたことが確認された。この二つの結果か

ら、RA に対する能力面での抑制要因が強いために両価的な状態にある人は、RA と元来好きな活動と組み合わせることにより「わくわくする」という直感的なポジティブ感情の高まりとともに動機づけられるという可能性が示された。また、RA に対する能力面がさほど抑制要因となっていない人にとっては、タスクの簡素化が動機づけとして有効である可能性が示された。

追跡調査の結果から、我々のメッセージ設計は実際に行動を変容するきっかけを提供できるものである可能性が示された。我々は、本調査時、最も動機づけられるメッセージの選択を指示したのみであり、その後行動するように教示はしなかった: 今後実行するように、とも、手元にメモするように、とも、覚えておくように、とも指示しなかった。それにも関わらず、彼らの一部は完全に主体的に、選んだメッセージの内容を手掛かりに行動を変えはじめた。このことは、心理的側面から行動の阻害要因を特定し、これに対処する具体的なヒントを提示することが、彼らの両価性の軽減と主体性の向上を促した可能性を示唆している。先行研究では行動計画を自ら立てることの有効性が確認されているため、その補助ツールとしても本研究の知見は役に立つかもしれない。

## 8. まとめ

生活習慣病の世界的な増加を踏まえると、健康的な行動の価値を認識しながらも先延ばしがちな人を動機付けることは極めて重要である。本研究では、文献レビューを通じてこうした両価的な人向けの 2 種類の動機づけプロファイルを導出し、LLM 生成を用いた動機づけメッセージを設

計した。これらのメッセージの有効性を評価するオンライン調査 (N = 548) の結果、両タイプとも参加者の動機づけに効果的であることが示された。さらに、知覚される能力が行動の抑制要因として強い参加者は「ポジティブ感情促進型」に、そうではない参加者は「ネガティブ感情低減型」に動機づけられる傾向が示された。今後、行動の実践率を高めるために、両価性に関するモチベーションプロファイルの精緻化を進めていく。

## 参考文献

- [1] Katharina Feil, Julian Fritsch, and Ryan E. Rhodes.: The intention-behaviour gap in physical activity: a systematic review and meta-analysis of the action control framework. *British Journal of Sports Medicine* 57, 1265 – 1271 (2023).
- [2] Piers Steel.: The nature of procrastination: a meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychological Bulletin* 133, 1, 65–94 (2007).
- [3] Nicholas W.S. et al.: The global burden of metabolic disease: Data from 2000 to 2019. *Cell Metabolism* 35, 3, 414–428 (2023).
- [4] Michelle Bak and Jessie Chin.: The potential and limitations of large language models in identification of the states of motivations for facilitating health behavior change. *Journal of the American Medical Informatics Association* 31, 9, 2047–2053 (2024).
- [5] Jeni Paay, Jesper Kjeldskov, Mikael B Skov, Lars Lichon, and Stephan Rasmussen.: Understanding individual differences for tailored smoking cessation apps. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 1699–1708 (2015).
- [6] Rosenstock, Irwin M.: The health belief model and preventive health behavior. *Health education monographs* 2.4: 354-386 (1974).
- [7] Bandura, Albert, and Richard H. Walters.: Social learning theory. Vol. 1. Englewood Cliffs, NJ: Prentice hall, (1977).
- [8] Prochaska, James O., and Carlo C. DiClemente.: Stages and processes of self-change of smoking: toward an integrative model of change. *Journal of consulting and clinical psychology* 51.3 (1983).
- [9] Hettema, Jennifer, Julie Steele, and William R. Miller.: Motivational interviewing. *Annu. Rev. Clin. Psychol.* 1.1. 91-111 (2005).
- [10] Michie, Susan et al.: The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Annals of behavioral medicine* : a publication of the Society of Behavioral Medicine vol. 46,1: 81-95 (2013).
- [11] Jisoo Lee, Erin Walker, Winslow Burleson, Matthew Kay, Matthew Buman, and Eric B Hekler.: Self-experimentation for behavior change: Design and formative evaluation of two approaches. In *Proceedings of the 2017 CHI conference on human factors in computing systems*. 6837–6849 (2017).
- [12] Susan Michie, Maartje M. van Stralen, and Robert West.: The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science* 6 (2011).
- [13] Evans, J. and Stanovich, K. E.: Dual-Process Theories of Higher Cognition Advancing the Debate, *Perspectives on Psychological Science*, Vol. 8, 3, pp. 223–241 (2013).
- [14] Fogg, B. J.: A behavior model for persuasive design, In *Proceedings of the 4th International Conference on Persuasive Technology (Persuasive '09)*, Article 40, pp. 1–7 (2009).
- [15] Fishbach, A. and Woolley, K.: The Structure of Intrinsic Motivation, *Annual Review of Organizational Psychology and Organizational Behavior*, Volume 9, pp. 339-363 (2022).
- [16] Milkman, K. L. et al.: Holding the Hunger Games Hostage at the Gym: An Evaluation of Temptation Bundling *Management Science*, Vol. 60, No. 2 (2013).
- [17] Kirgios, E. L. et al.: Teaching temptation bundling to boost exercise: A field experiment, *Organizational Behavior and Human Decision Processes*, Volume 161, Supplement, pp. 20-35 (2020).
- [18] Burrhus F Skinner.: Operant behavior. *American psychologist* 18, 8, 503 (1963).
- [19] World Health Organization.: Guidelines on physical activity and sedentary behaviour. (2020), <https://iris.who.int/handle/10665/336656> (参照 2025/11/28)
- [20] Getinet Ayano.: Dopamine: Receptors, Functions, Synthesis, Pathways, Locations and Mental Disorders: Review of Literatures. *Journal of Mental Disorders and Treatment* 2 (2016).
- [21] Koepf, M.J. et al.: Evidence for striatal dopamine release during a video game, *Nature*, Vol. 393, pp.266-268 (1998).
- [22] Takahashi, H. et al.: Dopamine D1 Receptors and Nonlinear Probability Weighting in Risky Choice, *Journal of Neuroscience*, Vol. 30, No. 49, pp. 16567-16572 (2010).
- [23] Takahashi, K. et al.: Imaging the passionate stage of romantic love by dopamine dynamics, *Frontiers in human neuroscience*, Vol. 9, 191 (2015).
- [24] OpenAI: GPT-4o, <https://openai.com/index/hello-gpt-4o/> (参照 2025/11/28).
- [25] Edward L Deci, Haleh Eghrari, Brian C Patrick, and Dean R Leone.: Facilitating internalization: The self-determination theory perspective. *Journal of personality* 62, 1, 119–142 (1994).
- [26] Terry Duncan Edward McAuley and Vance V. Tammem.: Psychometric Properties of the Intrinsic Motivation Inventory in a Competitive Sport Setting: A Confirmatory Factor Analysis. *Research Quarterly for Exercise and Sport* 60, 1, 48–58 (1989).
- [27] Richard M. Ryan, James P. Connell, and Robert W. Plant.: Emotions in nondirected text learning. *Learning and Individual Differences* 2, 1,1–17 (1990).
- [28] Richard M. Ryan, Richard Koestner, and Edward L. Deci.: Ego-involved persistence: When free-choice behavior is not intrinsically motivated. *Motivation and Emotion* 15, 185–205 (1991).

# 過去施策の有効性推定に基づくユーザー特性と行動を考慮した PUSH 配信手法

三浦優明 <sup>†1</sup> 吉川裕木子 <sup>†2</sup> 熊谷香織 <sup>†3</sup> 塚本航也 <sup>†4</sup> 山田拓也 <sup>†5</sup> 鈴木喬 <sup>†6</sup>  
株式会社 NTT ドコモ

## 1. 概要

本論文では、エコ行動アプリの利用促進を目的として、特性の異なる過去施策の結果を活用する新たな PUSH 配信手法を提案する。多くの PUSH 配信では、認知バイアス等の「ユーザー特性」や、アプリの利用履歴等の「ユーザー行動」を踏まえた文言が活用されている。しかし、「ユーザー特性」に基づく文言を繰り返し配信すると飽きや忌避感が生じたり、「ユーザー行動」に基づく配信では非アクティブユーザーには効果が及ばないという課題が存在した。本研究では、「ユーザー特性」と「ユーザー行動」の双方を考慮した PUSH 配信手法を提案する。具体的には、ユーザー特性やユーザー行動に基づく過去施策結果を用いて、各ユーザーへの施策有効性を推定し、その推定結果に基づいて配信文言を決定する手法である。本提案手法の有効性の検証のために、実際にユーザーに配信を行うことで比較検証を実施した。PUSH 開封率では本提案手法の効果は認められなかったが、PUSH 開封後のエコ行動記録回数の増加幅ではランダム配信群を上回った。さらに、各施策効果の推定結果を用いた分析により、施策特性の違いが文言の出し分け効果に与える影響についても検証した。

## 2. はじめに

### 2.1. 背景

近年、気候変動や資源枯渇といった地球規模の環境課題が広く注目されており、持続可能な社会の実現に向けた取り組みの重要性が高まっている。企業活動においても、環境へ配慮した事業活動は社会的責任として重要度が増しており、サステナビリティ経営の実践が求められている。このような背景から、NTT ドコモではサステナビリティ経営の一環として、ユーザーがエコ行動を記録するスマートフォ

A Push Notification Method Based on Estimation of Multiple Campaign Effects Using Machine Learning

<sup>†1</sup> MASAAKI MIURA, NTT DOCOMO, INC.

<sup>†2</sup> YUKIKO YOSHIKAWA, NTT DOCOMO, INC.

<sup>†3</sup> KAORI KUMAGAI, NTT DOCOMO, INC.

<sup>†4</sup> KOYA TSUKAMOTO, NTT DOCOMO, INC.

<sup>†5</sup> TAKUYA YAMADA, NTT DOCOMO, INC.

<sup>†6</sup> TAKASHI SUZUKI, NTT DOCOMO, INC.

ンアプリ「カボニューレコード」<sup>\*1</sup>を提供している。本アプリは、ユーザーが節電や節水、リサイクルなどの 19 種類のエコ行動を記録し、その行動による CO2 削減量を可視化できるものである。このアプリを通じて、ユーザーが日常生活の中で無理なく環境に配慮した行動を実践できる仕組みを提供している。

### 2.2. 課題と目的

カボニューレコードでは、ユーザーのエコ行動を促進するために PUSH 配信を活用しているが、その効果の向上が課題となっている。従来の PUSH 配信では、個々人にパーソナライズされた文言がユーザーの行動変容に有効であることが示されている [1, 2]。具体例としては、個々のユーザーが持つ価値観や認知バイアスに合わせて配信メッセージを出し分けする「ユーザー特性」に基づく施策 [1] や、ユーザーが過去に実施した行動へのフィードバックをメッセージに含めることで行動を促す「ユーザー行動」に基づく施策 [2] が実施してきた。しかし、「ユーザー特性」や「ユーザー行動」に基づいた配信は以下の課題を抱えている。

- (1) ユーザー特性に基づく配信の課題：特定の認知バイアスや価値観に訴えかける文言を繰り返し配信することで、ユーザーによってはその文言に対する飽きや忌避感を抱く可能性がある [3]。また、認知バイアスや価値観に基づく文言が効くかどうかは個々人の心理状況やタイミングに左右されることも指摘されている [4]。
- (2) ユーザー行動に基づく配信の課題：アプリの利用状況に応じたフィードバック文言を活用した配信では、非アクティブなユーザーには効果が及ばないことが課題である。

我々は、「ユーザー特性」と「ユーザー行動」の両方を考慮した最適なメッセージを出し分けることで、幅広いユーザー層に対して高い効果を発揮する PUSH 配信が実現できると考え、「ユーザー特性」と「ユーザー行動」の双方を考慮し

<sup>\*1</sup> 株式会社 NTT ドコモが提供するエコ行動記録アプリ「カボニューレコード」。URL: <https://cabonerecord.web.nttdocomo.ne.jp/top.html>

た新たな PUSH 配信手法を提案する。提案手法は、これまでに実施された「ユーザー特性」や「ユーザー行動」に関する過去施策の配信データを活用し、各ユーザーに対する施策の有効性を推定することで、より適切なメッセージを個別に出し分けるものである。また、本手法は NTT ドコモが提供するエコ行動記録アプリのユーザーを対象として実証を行い、その有効性を検証した。

### 3. 関連研究

従来、PUSH 通知の効果向上に向けては、配信タイミングの最適化や文言の工夫など、様々な改善が試みられてきた。その中でも近年注目されているのが、人間の思考の偏りを利用して行動を後押しする行動経済学的アプローチである「ナッジ (nudge)」である。ナッジは選択の自由を保持しつつ望ましい行動を促す手法として広く研究されており、PUSH 配信でも活用が進んでいる。例えば、健康管理アプリの PUSH 配信での検証 [5] やエコ行動記録アプリでの検証 [6] でナッジの有効性が確認されている。

PUSH 配信におけるナッジの一定の有効性が示される一方で、ナッジ文言への反応はユーザーの心理特性によって大きく異なることが指摘されている。このため、認知バイアス、価値観、性格特性 (Big Five) 等のユーザー特性に基づき、メッセージを出し分ける PUSH 配信手法が提案されている [1, 7]。文献 [1] ではユーザーの認知バイアスを推定し文言を出し分けており、文献 [7] では、ユーザーの性格特性 (BigFive) を推定し文言の出し分けを実施し、その有効性が確認された。しかし、ユーザー特性に基づく文言出し分けにはいくつか課題が存在する。1つ目は、同じ認知バイアスを持つユーザーに対してもバイアス文言への反応は異なることが指摘されている [7]。2つ目は、バイアス文言の有効性はユーザーの心理状況やタイミングによって変動する点である。同一ユーザーでも状況によってバイアス文言への反応が変わることが報告されている [4]。

ユーザーの過去行動や通知履歴を用いて PUSH 配信を最適化する手法も研究されている。文献 [8] では、健康管理アプリを対象に、ユーザーのアプリ利用履歴や通知反応に基づき、どの通知内容が行動を促しやすいかを分析し PUSH 配信を実施した。文献 [9] では、過去の通知反応ログを入力として、強化学習により「どのタイミングで通知を送るべきか」を最適化する手法が提案された。しかし、ユーザー行動を利用した PUSH 配信は、行動履歴が存在しない非アクティブユーザーにはアプローチできないという課題が存在する。

本研究では、過去施策の有効性推定モデルを活用し、「ユーザー特性」と「ユーザー行動」を考慮した PUSH 配信手法を提案することで、ユーザー行動に基づく PUSH 配信手法

では困難であった非アクティブユーザーへのアプローチを可能にするとともに、ユーザー特性のみに依存した場合に生じる文言効果のばらつきや状況依存性といった課題を軽減することを目指す。

## 4. 提案手法

ユーザー特性とユーザー行動の双方を用いて文言を出し分ける PUSH 配信手法について説明する。提案手法では、過去に実施された複数施策の配信結果をもとに、各ユーザーに対する施策有効性を推定し、その推定結果に基づいて最適な文言を選択することで文言の出し分けを行う。以降、(4.1) にて「ユーザー特性」と「ユーザー行動」に関する過去施策の内容を示し、(4.2) にて施策有効性の推定手法、文言の出し分け手法、(4.3) にて施策有効性の推定モデルの学習設定について説明する。

### 4.1. 提案手法で用いる過去施策

提案手法で活用する 3 つの過去施策内容と結果概要を紹介する。

#### 4.1.1 過去施策 1：フィードバック文言

提案手法で活用する 1 つ目の過去施策は、2024 年 12 月、2025 年 1 月に実施したフィードバック文言を活用した施策である。本施策は「ユーザー行動」に関する施策で、アプリ内でユーザーが過去に記録したエコ行動の内容に対し、その履歴を反映したフィードバック文言を配信するものであった。施策結果として、フィードバックなし文言を配信した群と比較して、フィードバック文言を配信した群の方が開封率およびエコ行動の増減数のいずれも高い傾向が確認された。

#### 4.1.2 過去施策 2：認知バイアスによる出し分け

提案手法で活用する 2 つ目の過去施策は、2025 年 2 月・3 月に実施したユーザーの認知バイアスを活用した施策である。認知バイアスとは人間の意思決定や判断に影響を与える思考の偏りであり、本施策では「向社会性」「利得」「同調効果」「損失」「時間選好」の 5 つの認知バイアスが使用された。本施策は「ユーザー特性」に関する施策で、各ユーザーがどの認知バイアスを最も強く持っているかを推定し、推定結果によって認知バイアスに特化した文言を出し分けで配信するものであった。施策結果として、認知バイアスに応じて文言を出し分けた群は、ランダムに文言を配信した群と比較して、開封率およびエコ行動増減数が高い値となった。

#### 4.1.3 過去施策 3：ポジティブ/ネガティブ文言

提案手法で活用する 3 つ目の過去施策は、2025 年 4 月に実施したポジティブ/ネガティブな文言を活用した施策である。本施策は「ユーザー特性」に関する施策で、文言のトー

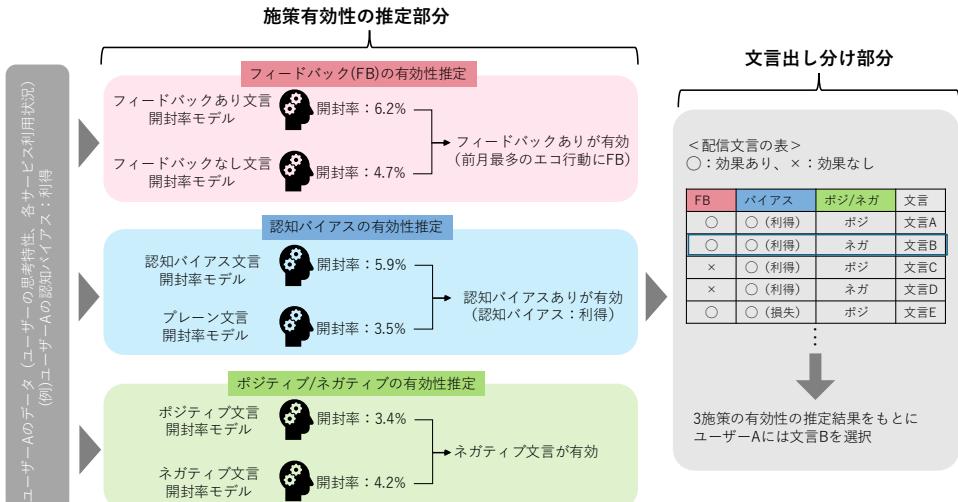


図 1 提案手法の概要

ンとしてポジティブ・ネガティブ・プレーンの3種類を用意し、配信対象者を3つのグループに分けてそれぞれの文言を配信した。施策結果としては、開封率とエコ行動増減数の両方において、プレーン文言を送った群が最も良い結果となった。ポジティブ文言とネガティブ文言を比較すると、ネガティブ文言を送った群の方が開封率が高く、エコ行動增加の増加幅も大きいことが確認された。

#### 4.2. 施策有効性の推定方法および文言出し分け方法

本論文では、3つの過去施策に対するユーザーごとの施策有効性を推定し、その推定結果に基づいて配信文言を出し分ける手法を提案する。提案手法の概要は図1に示す。施策有効性の推定は、各過去施策の配信データ（配信グループ分け、開封有無等）から開封率予測モデルを構築し、開封率の予測値が高い側を「有効」と判定する手順で実施した。具体的には、フィードバック文言の施策については、フィードバックあり文言とフィードバックなし文言それぞれについて開封率予測モデルを構築し、予測開封率が高い側を採用した。フィードバックありの文言が有効と判定された場合には、ユーザーが前月に最も多く記録したエコ行動に対するフィードバックを行うこととした。なお、前月のエコ行動を行っていないユーザーに対してはフィードバックをしないこととした。認知バイアスの施策については、認知バイアスありの文言と認知バイアスなしの文言（プレーン文言）のどちらがの予測開封率が高いかをもとに、認知バイアスあり/なしを決定した。認知バイアスありが有効と判断された場合には、ユーザーの持つ認知バイアスに対応した文言を配信する設定とした。ポジティブ/ネガティブ文言の施策については、ポジティブ文言とネガティブ文言の

それについて開封率予測モデルを構築し、予測開封率が高い側を選択した。以上の3施策の有効性推定の結果から、各ユーザーごとに3つの施策に関する有効性の組み合わせが決定された。配信文言は、この組み合わせに対応する文言を、表1に示すような配信用文言表から選択することで決定した。すなわち、「ユーザー特性」と「ユーザー行動」に基づく3つの過去施策ごとの有効性推定に基づいて、各ユーザーに最も適合すると判断される文言を配信することで、「ユーザー特性」と「ユーザー行動」の両方を考慮したPUSH配信を実現している。配信用文言は、大規模言語モデルであるGemini 2.5 Pro<sup>\*2</sup>に図2のようなプロンプトを入力することで作成した。

#### 4.3. モデル学習に利用したデータおよび設定

##### 4.3.1 学習に利用したデータ

本研究では、3つの施策それぞれの有効性推定モデルを構築するために、過去のPUSH配信ログとユーザー行動データを用いた。フィードバック施策の有効性推定には2024年12月・2025年1月の配信データ、認知バイアス施策の有効性推定には2025年2月・3月の配信データ、ポジティブ/ネガティブ施策の有効性推定には2025年4月の配信データを使用した。これらのデータには、配信時のグループ割り当ておよび配信後の開封有無が含まれており、開封フラグ(0:未開封, 1:開封)を目的変数としてモデルを学習した。また、特徴量としては、エコ行動記録アプリの利用履歴（前月・前週のエコ行動記録回数や最も多かったエコ行動種類

\*2 Googleが提供する大規模言語モデル。URL:<https://gemini.google.com/app>

施策 1 バイアス出し分け	施策 2 ポジ/ネガ	施策 3 フィードバック	PUSH タイトル	PUSH メッセージ
○ (向社会性)	ポジ	○ (節電)	あなたのエコ行動が地球を救っています！	最近の節電行動、素晴らしいです！地球の未来のために、今日もエコ行動を記録しませんか？
○ (向社会性)	ポジ	○ (節水)	あなたのエコ行動が地球を救っています！	最近の節水行動、素晴らしいです！地球の未来のために、今日もエコ行動を記録しませんか？
○ (向社会性)	ポジ	○ (フードロス)	あなたのエコ行動が地球を救っています！	最近のフードロス削減、素晴らしいです！地球の未来のために、今日もエコ行動を記録しませんか？
○ (向社会性)	ポジ	○ (リサイクル)	あなたのエコ行動が地球を救っています！	最近のリサイクル行動、素晴らしいです！地球の未来のために、今日もエコ行動を記録しませんか？
○ (向社会性)	ポジ	○ (その他)	あなたのエコ行動が地球を救っています！	最近のエコ行動、素晴らしいです！地球の未来のために、今日もエコ行動を記録しませんか？
○ (向社会性)	ポジ	×	あなたのエコ行動が地球を救っています！	あなたのエコ行動が、地球の未来を守る大きな力になります。今日もエコ行動を記録しませんか？
○ (向社会性)	ネガ	○ (節電)	地球の未来のために、今できることを.	最近の節電行動、素晴らしいです！地球温暖化や資源枯渇を防ぐためにもあなたのエコ行動が必要です！
○ (向社会性)	ネガ	○ (節水)	地球の未来のために、今できることを.	最近の節水行動、素晴らしいです！地球温暖化や資源枯渇を防ぐためにもあなたのエコ行動が必要です！
○ (向社会性)	ネガ	○ (フードロス)	地球の未来のために、今できることを.	最近のフードロス削減、素晴らしいです！地球温暖化や資源枯渇を防ぐためにもあなたのエコ行動が必要です！
○ (向社会性)	ネガ	○ (リサイクル)	地球の未来のために、今できることを.	最近のリサイクル行動、素晴らしいです！地球温暖化や資源枯渇を防ぐためにもあなたのエコ行動が必要です！
○ (向社会性)	ネガ	○ (その他)	地球の未来のために、今できることを.	最近のエコ行動、素晴らしいです！地球温暖化や資源枯渇を防ぐためにもあなたのエコ行動が必要です！
○ (向社会性)	ネガ	×	地球の未来のために、今できることを.	エコ行動を記録しましょう！地球温暖化や資源枯渇を防ぐためにもあなたのエコ行動が必要です！

表 1 9月配信で使用したメッセージの一部

等) やユーザーの思考特性 (認知バイアスや BigFive スコア), 許諾済みのユーザーから取得した各種サービスの利用状況を利用した.

#### 4.3.2 モデル構造と学習パラメータ

施策有効性の推定モデルには、決定木ベースの勾配ブースティングのフレームワークである LightGBM を用いた. LightGBM は大規模な表形式データに対して高い予測性能を持ち、学習が高速で行えることから、本研究の開封率予測モデルに採用した. 各推定モデルは、開封フラグ (0:未開封 / 1:開封) を目的変数とした二値分類モデルとして学習を行い、損失関数には二値分類に広く用いられる Binary logloss (クロスエントロピー誤差) を採用した. 学習の際に設定したパラメータは表 2 の通りである. それぞれの施策有効性の推定モデルについては、配信データの 80% を学習データ、20% を評価データとして利用した. なお、本研究におけるモデル学習は、各施策に対して 2 種類のパター

ン (フィードバックあり/なし、認知バイアスあり/ブレーン、ポジティブ/ネガティブ) の予測開封率を比較する必要があるため、各施策につき 2 種類のモデルを構築し、合計 6 つの開封率予測モデルを構築・学習した.

表 2 LightGBM の主な学習パラメータ

パラメータ名	設定値
objective	binary
metric	binary_logloss
boosting_type	gbdt
num_leaves	5
max_depth	3

ーク  
しをしない  
バンク利用  
スト利用  
トル持参  
ック持参  
・マイストロー持参  
分別

\*\*\*メッセージ本文 (50 文字程度) \*\*

図 2 文言作成用プロンプト例

## 5. 評価

### 5.1. エコ行動記録アプリ

NTT ドコモが提供するエコ行動記録アプリ「カボニューレコード」を対象として、提案する PUSH 配信手法の効果検証を実施した。本アプリは、ユーザーが日常のエコ行動を継続的に記録・可視化できるよう設計されており、その画面例を図 3 に示す。アプリでは、エアコンの設定温度の見直しやゴミの正しい分別など、全 19 種類のエコ行動を記録することができる。さらに、各エコ行動によって削減された CO2 排出量を日次で算出・表示する機能を備えており、ユーザーは自身の行動が環境負荷の低減にどの程度貢献したかを定量的に把握することが可能である。

また、アプリはユーザーに対して行動促進を目的とした PUSH 配信を行う機能を有しており、実際に配信される PUSH 通知の例を 4 に示す。本研究では、この PUSH 通知の受信を許可しているユーザーを対象として、提案手法の有効性を検証した。

### 5.2. 評価方法

本検証では、提案手法の有効性を検証するため、エコ行動記録アプリ「カボニューレコード」のユーザーを対象として A/B テストを実施した。評価対象は、PUSH 通知の受信を許可しているユーザーとし、2025 年 9 月 25 日に実施さ



図 4 アプリ PUSH メッセージ例

2

れたアプリ内 PUSH 配信を用いた。本実験では、対象ユーザーを以下の 2 群にランダムに割り当てた。

- **A 群 (モデル出し分け群)**：提案手法により、過去施策データから 3 施策それぞれの有効性を推定し、その推定結果の組み合わせに基づいてメッセージを配信した群。
- **B 群 (ランダム配信群)**：3 施策に対する有効性をランダムに選択してメッセージを配信した群。

性能評価では、最初に A 群と B 群の開封率を比較し、提案手法による文言出し分けがユーザーの開封行動に与える影響を確認した。さらに、PUSH 配信がユーザーのエコ行動に与える影響を評価するため、PUSH 配信を開封したユーザーを対象に、配信前 1 週間と配信後 1 週間におけるエコ行動の増減を算出し、A 群と B 群の比較を行った。追加分析として、施策特性の違いが文言の出し分けに与える影響についても検証した。具体的には、推定された施策効果について、認知バイアスの種類ごとや、ポジティブ・ネガティブのグループに分けて、開封率およびエコ行動増減数を分析した。

### 5.3. 配信結果

#### 5.3.1 開封率

A 群 (モデル出し分け群) と B 群 (ランダム配信群) の開封率を比較した結果を表 2 に示す。A 群の開封率は 2.46 % (開封数 342/配信者数 13,883)、B 群の開封率は 2.47 % (開封数 343/配信者数 13,861) であり、両群の開封率に大きな差は見られなかった。今回の検証では、推定された施策有

	開封率	開封数	配信者数
A 群 (モデル出し分け群)	2.46%	342	13883
B 群 (ランダム配信群)	2.47%	343	13861

表 3 開封率の比較

認知バイアス		開封率	開封数	配信者数
向社会性	A 群	5.63%	52	923
向社会性	B 群	5.30%	48	906
利得	A 群	8.44%	80	948
利得	B 群	8.14%	75	921
同調効果	A 群	4.68%	33	705
同調効果	B 群	5.23%	38	726
損失	A 群	4.38%	34	777
損失	B 群	4.07%	31	764
時間選好	A 群	5.37%	55	1025
時間選好	B 群	5.32%	53	997
プレーン	A 群	1.03%	88	9550
プレーン	B 群	1.03%	98	9505

表 4 認知バイアス別の開封率結果

ポジ/ネガ		開封率	開封数	配信者数
ポジティブ	A 群	0.88%	79	8982
ポジティブ	B 群	2.37%	163	6887
ネガティブ	A 群	5.37%	263	4901
ネガティブ	B 群	2.58%	180	6974

表 5 ポジティブ/ネガティブ別の開封率結果

効性に基づいてメッセージを出し分けた A 群と、ランダムにメッセージを配信した B 群との間で、開封率に明確な違いは確認されなかった。

施策の効果がユーザー特性によって異なる可能性を検討するため、認知バイアスの種類ごとにサブグループに分割して分析を行った。表 4 に認知バイアス別の開封率結果を示す。認知バイアスの種類ごとに結果を比較すると、向社会性や利得、損失の認知バイアスを持つユーザーに対しては、A 群の開封率が B 群を上回り、同調効果の認知バイアスを持つユーザーに対しては、B 群の開封率が A 群を上回る結果となった。この結果から、同調効果以外の認知バイアスでは A 群の開封率が B 群よりも高かったため、認知バイアス施策の有効性推定が開封率向上に寄与したと考えられる。

認知バイアス別の分析に続き、ポジティブ/ネガティブ文言の施策有効性が開封行動に与える影響を検討するため、ポジティブ文言とネガティブ文言に分けて分析を行った。その結果は表 5 に示すとおり、ポジティブ文言では B 群 (ランダム配信) の開封率が A 群 (モデル出し分け) を上回った一方、ネガティブ文言では A 群の開封率が B 群よりも高い値となった。この結果から、ネガティブ訴求のメッセージはユーザーごとの受容性の違いが大きく、施策有効性の推定結果による出し分けが有効な可能性が考えられる。

	開封者	エコ行動増減数	増減数の標準偏差
A 群	342	+2.33	9.40
B 群	343	+2.15	11.24

表 6 PUSH 開封者のエコ行動増減数の比較

認知バイアス		開封者	エコ行動増減数	増減数の標準偏差
向社会性	A 群	52	+0.87	11.52
向社会性	B 群	48	+2.83	11.17
利得	A 群	80	+3.62	10.09
利得	B 群	75	+2.97	10.63
同調効果	A 群	33	+2.30	8.69
同調効果	B 群	38	+0.79	11.44
損失	A 群	34	+4.68	9.94
損失	B 群	31	+1.26	11.79
時間選好	A 群	55	+3.22	8.46
時間選好	B 群	53	+0.04	16.53
プレーン	A 群	88	+0.56	7.63
プレーン	B 群	98	+3.14	7.35

表 7 認知バイアス別のエコ行動増減数

### 5.3.2 PUSH 開封者のエコ行動増減数

PUSH 通知を開封したユーザーを対象に、PUSH 配信前 1 週間と配信後 1 週間でエコ行動記録の増減数を比較した結果を表 6 に示す。A 群 (モデル出し分け群) の平均増減数は +2.33 回、B 群 (ランダム配信群) は +2.15 回となり、A 群の方がエコ行動の増減回数が大きい傾向が見られた。この結果は、施策有効性に基づく文言出し分けが PUSH 開封後の行動変容に一定の好影響を与える可能性を示しており、メッセージ内容の最適化がユーザーのエコ行動回数を増加させる要因となることが考えられる。

開封率の結果分析と同様に、認知バイアス別にエコ行動の平均増減数を算出した結果を表 7 に示す。利得、同調効果、損失、時間選好といったバイアスでは、A 群 (モデル出し分け群) の増減数がいずれも B 群 (ランダム群) を上回り、施策有効性の推定に基づく文言出し分けが、これらのバイアスに対して効果的に働く可能性が示された。一方、向社会性や損失の認知バイアスに関しては、A 群と B 群の間で大きな差は見られず、バイアスの種類によってメッセージ出し分けの効果が異なることが確認された。

続いて、メッセージのトーンがエコ行動記録の増減数に与える影響を検討するため、ポジティブ文言とネガティブ文言でグループ化して分析を行った。その結果は表 8 に示すとおり、ポジティブ文言では B 群 (ランダム配信群) の方が A 群 (出し分け群) よりも高い増減数を示した一方、ネガティブ文言では A 群の増減数が B 群を上回った。ネガティブ文言においては推定結果に基づいて文言を出し分けることで、ネガティブ方向のメッセージが特定ユーザーにより適合し、行動変容を促しやすくなる可能性が考えられる。ポジティブ文言はランダム配信である B 群の方がエコ行動増減数が大きく、ポジティブな文言は比較的幅広い

ポジ/ネガ	開封者	エコ行動増減数	増減数の標準偏差
ポジティブ A 群	79	+1.47	7.79
ポジティブ B 群	163	+1.93	11.23
ネガティブ A 群	263	+2.59	9.84
ネガティブ B 群	180	+2.35	11.27

表 8 ポジティブ/ネガティブ別のエコ行動増減数比較

ユーザーに受け入れられやすいため、施策有効性による出し分け効果が現れにくいことが考えられる。

## 6. 結論

本論文では、エコ行動記録アプリにおける PUSH メッセージの個別最適化を目的として、「ユーザー特性」と「ユーザー行動」に基づく過去施策の配信データから施策有効性を推定し、その推定結果に基づいてユーザーごとに文言を出し分ける手法を提案した。また、提案手法を実際のアプリ利用者に対して適用し、開封率およびエコ行動記録の増減数に対する効果を検証した。検証の結果、全体の開封率では A 群（モデル出し分け群）と B 群（ランダム群）の間に大きな差は見られなかったものの、認知バイアス別・ポジティブ/ネガティブ別といったサブグループに分割して分析した結果、特定のバイアス（利得・同調効果・時間選好）やネガティブ訴求の文言において、モデル出し分け群がランダム群より高いエコ行動増減数を示した。

本研究の結果から、特定の認知バイアスに基づく文言やネガティブ訴求のメッセージを送る際には、施策有効性の推定結果に基づいて文言を出し分けることで、エコ行動の記録増加につながる可能性があることが示された。これらの知見は、エコ行動アプリにおける PUSH 配信を「ユーザー特性」や「ユーザー行動」に応じてパーソナライズする取り組みを進めるうえで、有用な指針を与えるものである。

今後に向けて、以下のような課題が挙げられる。

- **配信タイミングの最適化**：ユーザーのアプリ利用状況をより詳細に把握し、開封しやすい時間帯に合わせて通知を配信することで、メッセージの受容性向上が期待される。例えば、利用頻度の高い時間帯やアクティブ状態にあるタイミングを検出し、それに応じて最適な配信スケジューリングを行うことが有効である。
- **ユーザーごとの推奨エコ行動の最適化**：ユーザーの行動履歴やアクティビティ度に応じて、推奨するエコ行動の内容を調整することが重要である。特に非アクティブユーザーに対しては、まず取り掛かりやすいエコ行動を提示するなど、ユーザー状態に応じた負荷の低い行動提案が行動促進につながると考えられる。

これらの課題に対して取り組みを進めることで、ユーザー

特性や行動履歴に基づいたより細かなメッセージのパーソナライズが可能となり、その結果として、エコ行動を一層促進する効果的な PUSH 配信の実現に寄与すると期待される。

## 参考文献

- [1] 吉川裕木子, 酒井亮勢, 鈴木明作, 鈴木 喬：認知バイアスと大規模言語モデルを活用したメッセージ最適化システムの提案, 第 113 回 モバイルコンピューティングと新社会システム (MBL) 研究会 (2024).
- [2] 塚本航也, 鈴木明作, 吉川裕木子, 勝間田優樹, 山田拓也, 石川太朗：エコ行動促進アプリにおける PUSH 配信を用いたフィードバック手法, AIoT 行動変容学会 第 9 回研究会 (BTI09) (2025).
- [3] Ma, H., Liu, X. and Shen, Z.: User fatigue in online news recommendation, *Proceedings of the 25th International Conference on World Wide Web*, International World Wide Web Conferences Steering Committee, pp. 1363–1372 (2016).
- [4] Kyung, N., Chan, J., Lim, S. and Lee, B.: Contextual Targeting in mHealth Apps: Harnessing Weather Information and Message Framing to Increase Physical Activity, *Information Systems Research*, Vol. 35, No. 3, pp. 1034–1051 (2023).
- [5] Valle, C. G., Nezami, B. T. and Tate, D. F.: Designing in-app messages to nudge behavior change: Lessons learned from a weight management app for young adults, *Organizational Behavior and Human Decision Processes*, Vol. 161, pp. 95–101 (online), <https://doi.org/10.1016/j.obhdp.2020.10.004> (2020). Creating Habit Formation for Behaviors.
- [6] Kim, H. B., Iwamatsu, T., ichiro Nishio, K., Komatsu, H., Mukai, T., Odate, Y. and Sasaki, M.: Field experiment of smartphone-based energy efficiency services for households: Impact of advice through push notifications, *Energy and Buildings*, Vol. 223, p. 110151 (online), <https://doi.org/10.1016/j.enbuild.2020.110151> (2020).
- [7] Janković, A., Kolenik, T. and Pejović, V.: Can Personalization Persuade? Study of Notification Adaptation in Mobile Behavior Change Intervention Application, *Behavioral Sciences (Basel)*, Vol. 12, No. 5, p. 116 (online), 10.3390/bs12050116 (2022).
- [8] Bidargaddi, N., Pituch, T., Maaieh, H., Short, C. and Strecher, V.: Predicting which type

- of push notification content motivates users to engage in a self-monitoring app, *Preventive Medicine Reports*, Vol. 11, pp. 267–273 (online), <https://doi.org/10.1016/j.pmedr.2018.07.004> (2018).
- [9] Yuan, Y., Muralidharan, A., Nandy, P., Cheng, M. and Prabhakar, P.: Offline Reinforcement Learning for Mobile Notifications (2022).

# Big Five 性格特性に基づく質問を用いた 内省誘発と決定遅延評価の試み

上野貴弘 <sup>†1</sup>  
福山大学      中野美奈 <sup>†2</sup>  
福山大学      中道上 <sup>†3</sup>  
福山大学/アンカーデザイン株式会社

## 1. はじめに

近年のウェアラブルデバイス普及に伴い、日常生活における情動や行動を生理指標と自己報告を組み合わせてトラッキングする Ecological Momentary Assessment (EMA) やデジタルフェノタイピングが多く研究されている [1, 2]。これらの手法の多くは、リッカート尺度や感情選択、スライダー尺度を基にしたインターフェースからユーザが最終的に選択した回答値とタイムスタンプのみを記録する。一方で、自己報告と生理指標の乖離が多く報告されており [3]、その要因の一つとして、回答に至るまでの迷いや判断の不確実性といった内省プロセスをトラッキング設計の中で扱っていないことが挙げられる。実際の回答行動では、「判断に自信がない」「どちらともいえない」といった状態がしばしば生じるが、従来のインターフェースはこうした不確実性を明示的に表明する手段を備えておらず、最終回答のみから回答過程を復元することは難しい。

そこで本研究では、「○」や「×」の明確な回答に曖昧回答オプション「?」を加えた自己報告インターフェースを設計し、これに決定遅延を組み合わせることで、回答過程そのものを行動指標として扱う内省センシングフレームワークを構築する。本論文では、このフレームワークの基盤的検証として、曖昧回答の利用と決定遅延の関係を評価する。質問内容としては、Big Five 性格特性に基づく 10 項目の質問を用いる。

Big Five 性格特性は、外向性、協調性、勤勉性、神経症傾向、開放性の 5 次元から成る性格モデル [4] であり、本研究ではこの枠組みに沿って質問項目を設計している。例えば、「あなたは、活発な人ですか?」「あなたは、控えめな人ですか?」(外向性)、「あなたは、人に気を遣う人間ですか?」「あなたは、わがままな人ですか?」(協調性)といった性格判断項目を、肯定 (○) と否定 (×) の対となる形で提

示する。このような項目は、回答者の中で自己特性について照会する必要がある場合に、一定の内省過程を伴う。また、肯定項目と否定項目をペアとして提示することで、「○」か「×」のどちらか一方のみが心理的に選択しづらいという状況を避け、両シンボルの心理的ハードルが極端に偏らないように設計している。

Big Five 性格特性に関する質問を用いる理由は、内省センシングの検証という観点から、短時間の実験で回答特性を観測できる点にある。性格特性はその場の気分と比べて短時間では変動しにくく、各項目に対する回答傾向は 1 セッション内で安定している。そのため、長期間の EMA を実施せずとも、单一セッションの実験で、曖昧回答「?」の使用頻度や決定遅延の分布といった指標を用いて、インターフェース設計や内省指標としての妥当性を評価できる。また、「○」と「×」の心理的ウェイトが揃えられているため、回答シンボル間の差は主として「内省の仕方」や「迷いの強さ」に起因するとみなせる。

さらに本研究では、異なる社会的状況における回答行動の比較も目的とする。メンタルヘルストラッキングは本来、個人が日常環境で単独に利用することを想定しているが、回答行動が他者の存在や場の構造といった社会的文脈の影響を受ける可能性は大きい。本実験では、監督者と対面で回答する実験室環境と、多数の学生が分散して着席する講義室環境の 2 条件を設定し、曖昧回答「?」の選択傾向と決定遅延が環境によってどのように変化するかを比較する。

以上を踏まえ、本研究の目的は次の 2 点である。

1. Big Five 性格特性に基づく質問において、曖昧回答「?」と決定遅延がどのように反映されるかを明らかにする。
2. 2 つの環境下 (実験室と講義室) での応答特性の変化を検証する。

この目的のために、曖昧回答オプションを含む自己報告インターフェースと決定遅延計測を組み合わせた実験を 2 つの環境 (実験室と講義室) で実施し、回答シンボルごとの回答率および決定遅延の統計量を比較・分析する。

Inducing Reflection and Evaluating Decision Latency through Questions Based on the Big Five Personality Traits

<sup>†1</sup> TAKAHIRO UENO, Fukuyama University

<sup>†2</sup> MINA NAKANO, Fukuyama University

<sup>†3</sup> NOBORU NAKAMICHI, Fukuyama University/ANCR DESIGN Inc.

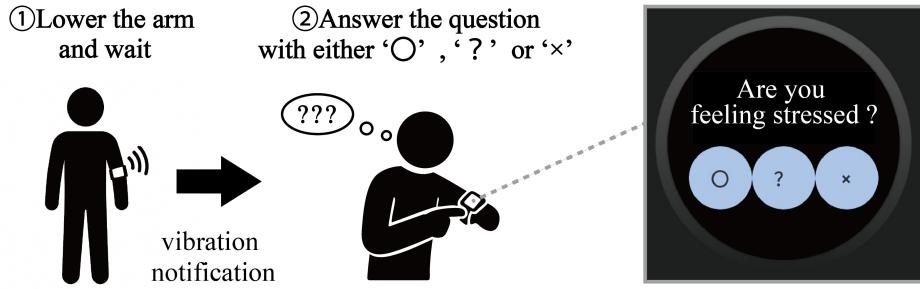


図1 曖昧回答オプション「?」と決定遅延を組み合わせた内省センシングフレームワーク

Figure 1 Introspective sensing framework combining the uncertain option '??' with decision-latency

## 2. 内省センシングフレームワークと質問設計

### 2.1. 内省センシングフレームワーク

図??に、本研究における内省センシングフレームワークの概要を示す。本フレームワークでは、メンタルヘルスに関する質問（例：「あなたはストレスを感じていますか？」）に対して、従来の「○」（はい）と「×」（いいえ）に加えて、明示的な曖昧返答オプション「?」を提示する。手順は、(1) 腕を下げて待機し、(2) 振動通知後に手首を上げて質問に回答する、という2段階から構成される。この過程で、応答時の手首の静止時間を加速度センサで計測し、振動提示から応答完了までの決定遅延（decision latency）として記録する。スマートウォッチを用いることで、時計で現在時刻を確認する自然な単純動作の中に内省を埋め込み、そのプロセスを行動信号として取得できる点が特徴である。

従来の Ecological Momentary Assessment (EMA) や digital phenotyping では、ユーザが最終的に選択した回答値とその時刻を主なデータとして扱い、自己報告と生理・行動指標との対応づけにより日常的な心理状態の推定が試みられてきた。しかし、これらの手法は「結果としての回答」のみを取得しており、回答に至るまでの迷いや判断の不確実性といった内省過程は構造的に捉えられていない。特に、気分・感情に関する質問では、その場の感情状態や社会的望ましさの影響から、ポジティブな方向への回答に偏ることが知られており [5, 6]、こうしたバイアスが自己報告と生理指標の乖離を生む一因となる。

図1に示すように、本フレームワークでは、明示的な「?」オプションにより、参加者が「判断できない」「自信がない」といった不確実な状態を回答として表明できるようにするとともに、加速度センサにより応答までの決定遅延を計測することで、内省の深さや判断困難さを反映した行動指標を得ることを目指している。ただし、このフレームワークの有効性を検証するためには、回答に内省過程が生じやすく、かつ気分・感情質問で問題となるポジティブバイアスの影響を受けにくい質問内容を用いる必要がある。次節では、この目的のためにビッグファイブ性格特性に基づく質問を

表1 Big Five 性格特性に基づく質問項目

Table 1 Personality questions based on Big Five framework

性格特性	質問
外向性	あなたは、活発な人ですか？
外向性	あなたは、控えめな人ですか？
協調性	あなたは、わがままな人ですか？
協調性	あなたは、人に気を遣う人間ですか？
勤勉性	あなたは、しっかりものですか？
勤勉性	あなたは、だらしない人ですか？
神経症傾向	あなたは、冷静な性格ですか？
神経症傾向	あなたは、心配性ですか？
開放性	あなたは、新しいことが好きですか？
開放性	あなたは、アイデアを出すことは苦手ですか？

用いる理由について述べる。

### 2.2. Big Five 性格特性に基づいた質問による内省誘発

前節で述べたように、気分・感情の直接質問は、その場の感情状態や社会的望ましさの影響からポジティブな方向への回答に偏りやすく、内省過程に伴う不確実性や迷いを多様な応答パターンとして捉えるには適していない。そこで本研究では、質問内容として、Big Five 性格特性（外向性、協調性、勤勉性、神経症傾向、開放性）の枠組み [4] に基づく 10 項目の性格特性質問を採用した。Big Five は、人格記述語の共分散構造から抽出された 5 因子モデルであり、比較的長期に安定した性格傾向を記述する尺度として用いられている。

表1に、本実験で用いた質問項目を示す。これらの質問はいずれも、参加者自身の性格傾向についての自己評価を求めるものであり、「今の気分」や一時的な感情状態を問う項目とは異なり、即座に肯定・否定しにくい内容である。性格特性に関する判断は本質的に抽象度が高く、状況や文脈によって揺らぎうるため、参加者は「自分は本当にそうと言

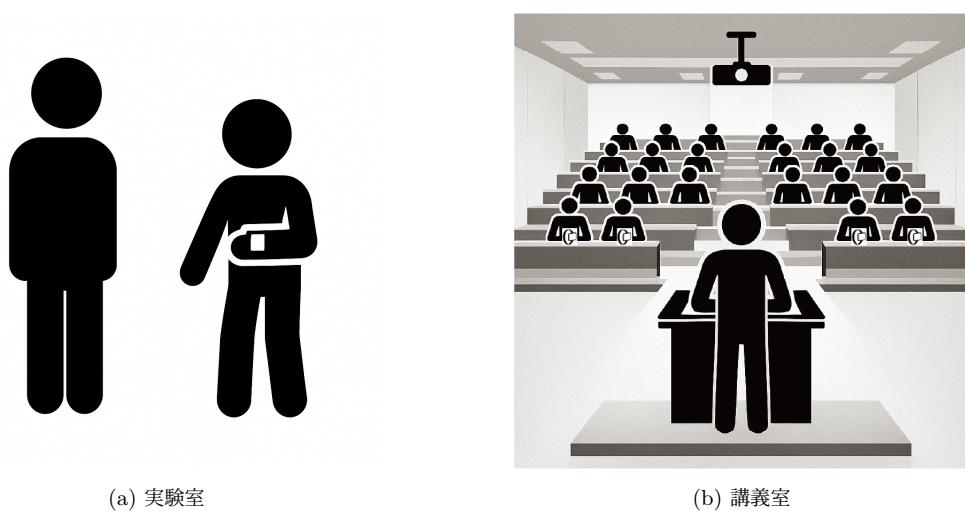


図 2 実験環境

Figure 2 Experimental settings

えるのか」「場合によっては当てはまらないのではないか」といった自己認識について一定の熟考を要する。そのため、回答に至るまでの内省過程が自然に生じやすく、前節で示した内省センシングフレームワークにおける曖昧回答や決定遅延の分析に適していると考えられる。

さらに、各性格特性については、肯定的なフレーミングに基づく質問と否定的なフレーミングに基づく質問を1項目ずつ配置し、ペア構造を持たせている。例えば、外向性に関しては「あなたは、活発な人ですか?」と「あなたは、控えめな人ですか?」、協調性に関しては「あなたは、わがままな人ですか?」と「あなたは、人に気を遣う人間ですか?」といった対になった項目を用いている。性格特性を明確に自認している参加者であれば、これらのペアに対して一方には「○」、他方には×を選択することが期待される一方で、自己認識が曖昧な場合や文脈に依存する場合には、どちらの項目に対しても判断に迷いが生じる可能性が高い。このペア構造により、肯定・否定のいずれか一方のみが「心理的に正しい」選択として過度に優位になることを避け、各シンボル(「○」「?」「×」)の心理的ハードルを相対的に均質化することを意図している。

このように、ビッグファイブ性格特性に基づく質問項目は、一時的な感情状態に左右されにくく、自己概念との照合を伴う内省過程を自然に誘発する設計になっている。本研究では、3章で述べるように、これら10項目に対するスマートウォッチ上での回答において、明示的な曖昧選択肢(?)と決定遅延を同時に計測することで、内省センシングフレームワークの基礎的特性を実験的に検証する。

### 3. 実験

本研究では、図2に示すように、実験実施者と参加者が対面する実験室環境と、多数の学生が同時に回答する講義室環境の2条件で実験を実施した。実験室環境では、実験実施者立会いのもと参加者と1対1という状況で、スマートウォッチに提示される質問に回答させた。講義室環境では、大教室において一斉実験の形式で実施した。学生は段々状の机・椅子に分散して着席した状態で、各自の席からスマートウォッチを用いて回答した。デバイスの配布は、最前列の学生から順に行い、実験が終了した参加者は、近くの後列の学生にスマートウォッチを手渡すことで後方へ回していく方式とした。実験室条件の参加者は41名の健康な日本人(平均年齢 = 28.5 ± 13.0歳、男性36名、女性5名)である。講義室条件の参加者は91名(平均年齢 = 19.25 ± 1.1歳、男性85名、女性6名)である。本研究は福山大学研究安全倫理委員会の承認を得ており、参加前に口頭説明およびスマートウォッチ上での同意入力によりインフォームドコンセントを取得した。実験はいずれも屋内で実施した。

#### 3.1. 実験手順

図3に実験手順のシーケンス図を示す。参加者は次の試行を待つ間、腕を下げた状態を保つよう指示された。数秒の待機時間後、スマートウォッチが振動通知を発した。その後、参加者は時刻を確認するように手首を上げ、ディスプレイを見ることで質問を視認する。ディスプレイには1つの性格関連質問と3つの返答シンボル(「○」: はい、「?」: わからない、「×」: いいえ)が表示された。「?」が選択された場合、その選択が「○」と「×」のどちらに寄りであるかを回答し、「?○」または「?×」として記録された。統計分析にお

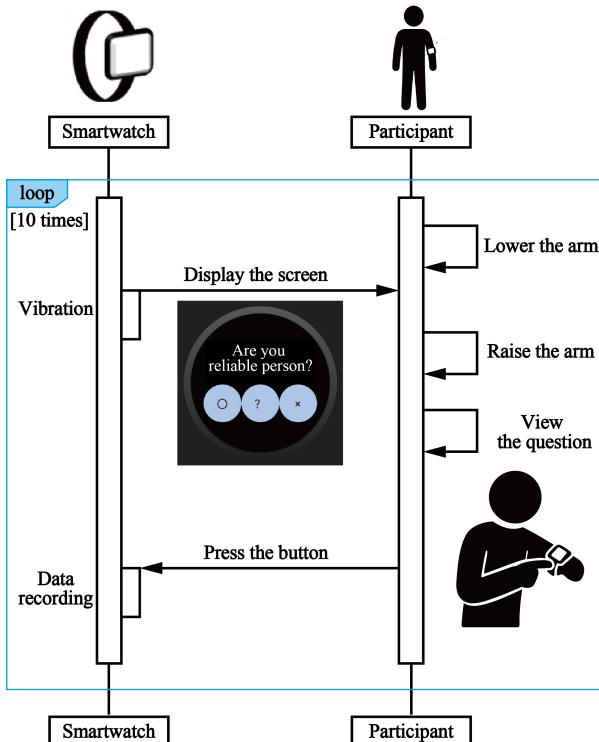


図3 実験手順のシーケンス図

Figure 3 Sequence diagram of experimental procedure

いでは、「?」と記された返答は「?○」と「?×」の両方のサブカテゴリーを包含する。なお、「?○」と「?×」への分類は実験分析目的のためであり、実際のメンタルヘルストラッキングシステムでは使用されない。参加者の一部は、「?○」と「?×」への追加質問なし（「○」、「×」、「?」の3つの回答オプションのみ）を利用した。

### 3.2. 決定遅延の計測について

決定遅延の計測には Google Pixel Watch 3 を使用し、約 16 Hz のサンプリング周波数で加速度計データを収集した。図 4 に「?」返答時の加速度センサの時系列データ例を示す。Question response time は、質問開始のためにスマートウォッチが振動してから参加者がシンボルを返答するまでの時間帯である。図 4 によると、しばらくすると加速度センサの値が増加しはじめ手首をあげる動作を開始している。また、「?」返答直前にはセンサの値が 0 に近い値が続くことで、手首が静止していることが分かる。これを決定遅延と定義し、応答の一連の動作を基に、Question response time における加速度センサの時系列データを下記の 2 つの異なるフェーズに分けた。

1. 振動後の手首を上げる動作
2. 返答選択中の静止

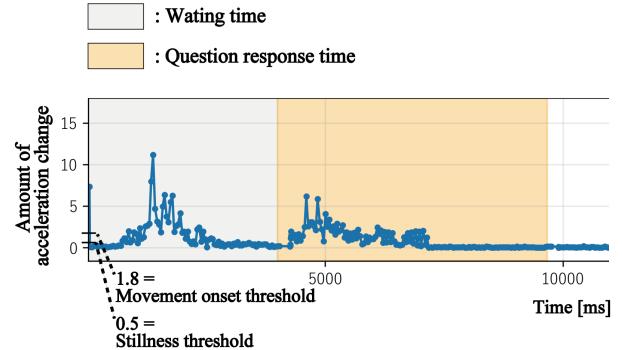


図4 「?」返答時の加速度センサの時系列データ

Figure 4 Time-series accelerometer data during '?' response

加速度計データから、次式で変化量 ( $\Delta_t$ ) を計算した:

$$\Delta_t = \sqrt{(x_t - x_{t-1})^2 + (y_t - y_{t-1})^2 + (z_t - z_{t-1})^2} \quad (1)$$

ここで、 $x_t, y_t, z_t$  は時間  $t$  における加速度 (m/s<sup>2</sup> 単位) である。決定遅延は、振動通知から返答選択までの間に発生する静止間隔 ( $\Delta_t < 0.5 \text{ m/s}^2$ ) の累積持続時間として定義された。動作開始しきい値 ( $1.8 \text{ m/s}^2$ ) と静止しきい値 ( $0.5 \text{ m/s}^2$ ) を適用して、動作と静止を自動的に区別した。また、 $K$  を単一試行で検出された静止間隔の総数とする。静止しきい値を下回る  $m_i$  個の連続サンプルからなる各静止間隔  $i$  の持続時間は次式で表した:

$$d_i = m_i \times (1/f) \quad (2)$$

ここで、 $f$  はサンプリング周波数である。試行における決定遅延  $DL$  は、すべての静止間隔の持続時間の合計として定義した:

$$DL = \sum_{i=1}^K d_i \quad (3)$$

また、各参加者の回答シンボル別の平均決定遅延  $\overline{DL}$  を計算する。これは、決定遅延が応答スタイルの違いに個人差があるためである。参加者  $j$  がシンボル  $s \in \{\circ, \times, ?, ?\circ, ?\times\}$  を選択した全試行における平均決定遅延  $\overline{DL}$  を次式で定義した:

$$\overline{DL}_{j,s} = \frac{1}{n_{j,s}} \sum_{i=1}^{n_{j,s}} DL_{j,s,i} \quad (4)$$

ここで、 $DL_{j,s,i}$  は参加者  $j$  がシンボル  $s$  を選択した  $i$  番目の試行における決定遅延であり、 $n_{j,s}$  はそのような試行の総数である。

### 3.3. 分析方法

Big Five 性格特性に基づく質問における、曖昧回答「?」と決定遅延の特性を分析するため、回答シンボルごとの使

表 3 回答シンボルごとの  $\overline{DL}$  の統計結果 (Mean±SD)Table 3  $\overline{DL}$  statistics by response type and symbol

	○	?	×	?○	?×
実験室 $\overline{DL}$	1295 ± 798	2352 ± 1973	1680 ± 1035	1663 ± 1074	2783 ± 2899
	41	25	40	13	13
講義室 $\overline{DL}$	1291 ± 3064	3177 ± 3580	1125 ± 1304	3536 ± 4476	2596 ± 2569
	91	3580	66	61	48

Note:  $N$  = Number of participants who used the symbol

用頻度と決定遅延の双方を分析対象とした。まず、各参加者について 10 間中の「○」、「?」、「×」の出現回数を集計し、環境ごとに平均値、標準偏差、中央値、最小値、最大値を算出した。これにより、「○」と「×」のバランスや、「?」使用のばらつきといった回答傾向を記述統計として整理する。あわせて、曖昧回答の内訳である「?○」「?×」についても、同様に出現回数および決定遅延を集計し、参考情報として整理した。決定遅延については  $\overline{DL}$  を用いた。シンボル間の決定遅延における有意差を検証するために、実験室条件と講義室条件それぞれについて、「○」「?」「×」の 3 水準を被験者内要因とするフリードマン検定を適用した。また、各回答シンボルのペアについて、ウィルコクソン符号付き順位検定を行い、2 シンボル間での決定遅延の有意差を検証した。各検定の有意水準は  $p < 0.05$  とした。

2 つの環境下（実験室と講義室）での応答特性の変化を検証するために、上記の記述統計およびシンボル間比較を両環境で同じ手順で行い、パターンの特性を比較した。

## 4. 結果と考察

### 4.1. Big Five 質問における曖昧回答と決定遅延

表 2 参加者のシンボルごとの回答頻度（全 10 間）

Table 2 Descriptive statistics of the number of responses for each symbol per participant (10 questions)

	Mean±SD	Max	Min	Median
実験室 ○	5.66 ± 1.39	9	3	6
	1.66 ± 1.76	7	0	1
	2.68 ± 1.52	6	0	3
講義室 ○	5.45 ± 1.52	9	2	5
	1.99 ± 1.73	6	0	2
	2.56 ± 1.48	6	0	2

表 2 に、各参加者が 10 項目に対して選択したシンボル（「○」「?」「×」）の出現数を示す。多くの参加者で「○」と「×」はいずれも複数回用いられており、10 項目中、「○」は概ね 3~7 回、「×」は 2~5 回の範囲で分布していた。一方

表 4 実験環境ごとの  $\overline{DL}$  [ms] の検定結果Table 4 Statistical testing results of  $\overline{DL}$  [ms] by experimental environment

	$p$ 値	$N$
実験室 ○, ×, ? Groups	<b>.004</b>	24
	<b>.037</b>	40
	<b>.002</b>	24
	.128	25
講義室 ○, ×, ? Groups	< <b>.001</b>	62
	<b>.042</b>	87
	< <b>.001</b>	62
	< <b>.001</b>	66

で、平均値・中央値ともに「○」の出現数は「×」より一貫して多く、性格特性に関する自己報告では、肯定的自己像を選好するバイアスや、「○」の方が心理的ハードルの低い応答として採用されやすい傾向が残存していると考えられる。ただし、ほとんどの参加者が「○」「×」の双方を複数回用いており、いずれか一方のみを選択し続ける極端なパターンは見られなかったことから、本研究の質問設計は、「○」と「×」のどちらか一方に過度に偏る設計にはなっていないと言える。一方、「?」の出現数については個人差が大きく、10 項目中 0 回であった参加者から、複数の項目で「?」を選択した参加者まで幅広い分布が見られた。中央値は少數回にとどまる一方で、最大値は複数項目に達しており、「?」を積極的に用いる参加者群が存在することが示される。

次に、回答シンボルごとの決定遅延 ( $\overline{DL}$ ) の違いを検討した結果を表 3 に示す。実験室条件・講義室条件のいずれにおいても、「○」および「×」選択時の  $\overline{DL}$  は相対的に短く、「?」選択時の  $DL$  が最も長い値を示した。また、 $\overline{DL}$  にフリードマン検定およびウィルコクソン符号順位検定を適用結果を表 4 に示す。フリードマン検定により 3 シンボル間の  $\overline{DL}$  に有意差が認められ、ペア比較の結果、「○」は「×」および「?」より有意に短く、「?」は少なくとも一方のシンボルより長い  $\overline{DL}$  を示した。

以上より、Big Five 性格特性に基づく性格判断課題にお

いて、肯定的回答「○」は頻度が高く決定遅延も短いことから、他のシンボルに比べ心理的ハードルの低い応答として機能していると解釈できる。一方、曖昧回答「?」は、記述統計および統計検定の両面から、他のシンボルより一貫して長い決定遅延を伴う応答として現れており、判断に迷う項目や自己認識が揺らぐ項目における内省の深さ・判断困難さを反映した指標となりうることが示唆された。

#### 4.2. 実験室と講義室における環境差

表5 各シンボルの回答割合

Table 5 Proportion of participants who used each response symbol

	○	?	×
実験室 N	41	25	40
回答者割合	100%	61%	98%
講義室 N	91	66	87
回答者割合	100%	73%	96%

実験室環境と講義室環境におけるシンボル使用状況を比較した結果を表5に示す。両環境とも、全参加者が少なくとも1回は「○」を選択し、「×」もほぼ全員が使用していた。一方、「?」を少なくとも1回使用した参加者の割合は、実験室条件で61%、講義室条件で73%と、講義室の方12%が高かった。

また、表2から、両環境とも1人あたりの「○」の回答数は平均5~6件、「×」は2~3件程度であり、全体として「○」「×」のバランスは大きく崩れていないことが確認された。一方、「?」の使用頻度には環境にかかわらず大きな個人差があり、一度も利用しない参加者から複数項目で「?」を選択する参加者まで多様なパターンが存在した。

決定遅延に関しては、両環境とも「?」選択時の $\overline{DL}$ が「○」選択時より一貫して長かったが、講義室条件ではさらに「×」よりも「?」が有意に長い $\overline{DL}$ を示し、「?」が3シンボルの中で最長となる傾向がより明確であった(表3, 4)。このことから、「?」の選択は単に選択肢が1つ増えたことによる偶発的な選択ではなく、他のシンボルより長い静止時間を伴う「迷い」や「判断困難さ」と結びついている可能性が高いといえる。

さらに、講義室環境では「?」使用者の割合が増加し、かつ「?」選択時の $\overline{DL}$ も長くなる傾向が見られたことから、監督者との対面がない状態や他者と同席している状況では、曖昧さや判断の迷いをより表しやすい可能性が示唆される。また、メンタルヘルストラッキングの文脈では、監督者のいない状況下での回答が主であるため、この結果から講義室環境が日常生活でのトラッキング条件に近い可能性が考えられる。

#### 5. おわりに

本研究では、スマートウォッチ上に「○」「?」「×」の回答選択肢を用意する自己報告インターフェースに加え、決定遅延を同時に計測する内省センシングフレームワークを構築した。そして、Big Five 性格特性に基づく性格質問を用いてその基礎的特性を検証した。結果として、肯定的回答「○」は他のシンボルと比べて使用頻度が高く決定遅延も短い一方で、曖昧回答「?」は一貫して最も長い決定遅延を伴うことが示され、「判断がつかない」状態や自己認識の揺らぎを反映した応答として機能している可能性が示唆された。また、実験室と講義室という異なる社会的文脈の比較から、講義室環境では「?」を用いる参加者の割合が高まり、「?」選択時の決定遅延もより長くなる傾向が確認され、曖昧さや迷いの表明が環境条件によって変化しうることが示された。今後は、「今の気分はどうですか?」などの実運用を想定した質問内容への拡張や、心拍変動などの生理指標との組み合わせによるメンタルヘルストラッキングへの応用を進めることで、本フレームワークの有効性と限界をより詳細に検証していきたい。

**謝辞** 本研究はJSPS科研費JP25K21258の助成を受けたものです。

#### 参考文献

- [1] Arvey, M. F., Schatten, H. T., Haradhvala, N. and Miller, I. W.: Ecological momentary assessment (EMA) of depression-related phenomena, *Current opinion in psychology*, Vol. 4, pp. 21–25 (2015).
- [2] Onnela, J.-P. and Rauch, S. L.: Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health, *Neuropsychopharmacology*, Vol. 41, No. 7, pp. 1691–1696 (2016).
- [3] Siepe, B. S., Tutunji, R., Rieble, C. L., Proppert, R. K. and Fried, E. I.: Associations between ecological momentary assessment and passive sensor data in a large student sample., *Journal of Psychopathology and Clinical Science* (2025).
- [4] John, O. P., Srivastava, S. et al.: The Big-Five trait taxonomy: History, measurement, and theoretical perspectives, University of California Berkeley (1999).
- [5] Caputo, A.: Social desirability bias in self-reported well-being measures: Evidence from an online survey, *Universitas Psychologica*, Vol. 16, No. 2, pp. 245–255 (2017).
- [6] Krumpal, I.: Determinants of social desirability bias in sensitive surveys: a literature review, *Quality & quantity*, Vol. 47, No. 4, pp. 2025–2047 (2013).

# ドライバの運転行動変容による 自動車からの排出ガス低減に関する研究

池本雅里<sup>†1</sup> 守山恵太<sup>†1</sup> 今井幸平<sup>†1</sup> 峯尾健太郎<sup>†1</sup> 塚越崇博<sup>†1</sup> 西岡寛真<sup>†1</sup> 川口文悟<sup>†1</sup>  
トヨタ自動車<sup>†1</sup>

荒川豊<sup>†2</sup>  
九州大学<sup>†2</sup>

## 1. はじめに

近年、環境意識に対する社会的関心が高まっており、自動車は環境負荷に大きな影響を及ぼすため、持続可能な技術の開発が求められている。トヨタ自動車ではカーボンニュートラルの実現に向けて、電気自動車、ハイブリッド車、プラグインハイブリッド車など、複数の技術や手段を並行して開発する「マルチパスウェイ戦略」を推進している。その中で、エネルギー供給やインフラの観点から、内燃機関は引き続き重要な役割を担うと考えられている[1]。

内燃機関は有害なガスを排出するが、自動車からの排出ガス規制は図1に示すように年々厳しくなっており、近年ではRDE(Real Driving Emissions)規制による実路走行に対する規制や、車載センサを搭載して実路走行時の排出ガスを計測するOBM(On-Board Monitoring)規制[2]など、従来の台上モード走行での排出ガス規制に加えて、実際の走行環境における排出ガスへの注目が高まっている。

筆者らは、内燃機関の「リアルワールドゼロエミッション」化によってカーボンニュートラルの実現に貢献することを目指している。そのため、実路での排出ガス低減を目指し、交通流シミュレーションとドライビングシミュレータを連携させたデジタルツイン開発環境を構築し、実路走行時の排出ガス低減に向けたパワートレーン・排気システム制御による対策技術の開発に取り組んでいる[3][4]。

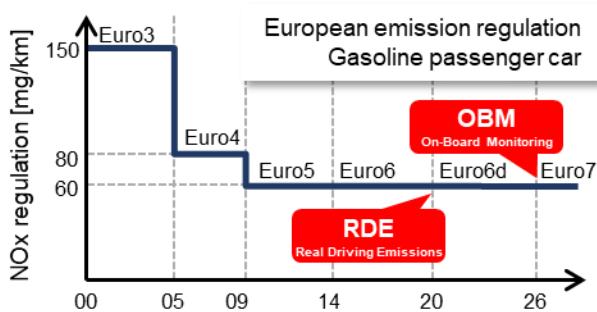


図1 欧州乗用ガソリン車のNOx排出規制値の推移

一般的に、排出ガス低減技術の開発では、エンジン単体での評価や、車輪をローラ上で回転させて走行状態を再現するシャシダイナモーティによるパターン走行試験が用いられる。表1に示すように、シャシダイナモーティでは、気温・風・路面などの外部条件の影響を排除し、決められたパターンで走行するため、ドライバによる差は小さく、排出ガスや燃費を精密かつ高い再現性で測定できるというメリットがある。一方、実路走行では、自転車や歩行者を含む交通状況によってドライバの運転操作が影響を受けるため、排出ガスはドライバの運転特性に大きく左右される。

本プロジェクトでは、「クルマ・ドライバ・交通流の三位一体で環境に優しい社会を実現する」ことを目標とし、従来から取り組んでいるパワートレーンや排気システムによる排出ガス低減技術の開発に加え、ドライバの運転操作や交通流システムに着目し、「リアルワールドゼロエミッション」と「Fun to Drive」の両立する技術を開発している。

特に、実路での排出ガスは交通状況に応じたドライバの運転操作の影響が大きいため、本研究では、ドライバの運転行動変容を促進することによる実路走行時の排出ガス低減ポテンシャルを評価した。

表1 シャシダイナモーティと実路走行の比較

	シャシダイナモーティ	実路走行
試験方法		
環境	気温・気圧の条件固定 路面振動も少ない	気象条件は成り行き 夜間走行の場合もあり 路面振動も多い
走り方	固定サイクル 車速追従の単独走行	交通流に沿った 前車追従の混走 自転車/歩行者の影響あり
ドライバ	プロのドライバ 運転特性バラつき小	一般のお客様 運転特性の影響大
排出ガス計測	据置型の分析計 (燃費・排出ガス精度 良)	車載排出ガス分析装置

†1 MASATO IKEMOTO, KEITA MORIYAMA, KOHEI IMAI, KENTATO MINEO, TAKAHIRO TSUKAGOSHI, HIROMASA NISHIOKA, BUNGO KAWAGUCHI, Toyota Motor Corporation

†2 YUTAKA ARAKAWA, Kyushu University

## 2. ドライバの運転特性の影響

### 2.1 ドライバによる排出ガスの差

研究を進めるにあたり、実路におけるドライバによる排出ガスの差を調査した先行研究[3]の結果を図2および図3に示す。試験には、車載排出ガス計測装置を搭載した2.0Lハイブリッドシステムの車両を用い、4名のドライバが名古屋空港線約2kmの区間を走行した。

計測対象の排出ガス成分はNOxおよびCOであり、その排出量を図2に示す。4名のドライバ間でNOxおよびCOの排出量に大きな差が認められた。また、図3には地図上にNOxの排出量をプロットした。ドライバによらず交差点付近に排出ポイントが集中する傾向はみられるものの、排出量や排出箇所にはドライバごとの差異が確認された。

今回の走行区間は片側3車線で交通量が多く、首都高速の入口・出口を含む複雑な交通流を有しているため、他車の挙動の影響を受けやすく、走行時の周囲の交通流によってドライバごとのアクセル・ブレーキ操作の特徴が現れ易い環境であったと考えられる。その結果として、排出ガスの排出量や排出箇所に差が生じたものと推察される。

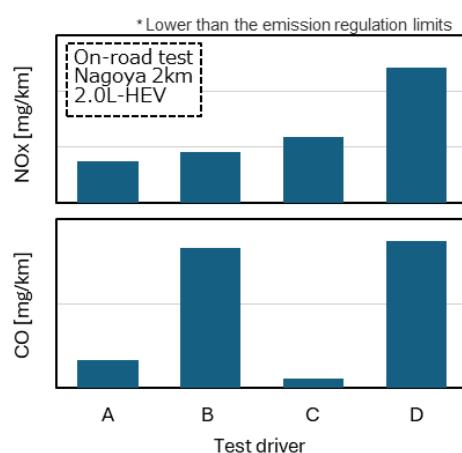


図2 ドライバごとの NOx・CO 排出量の比較

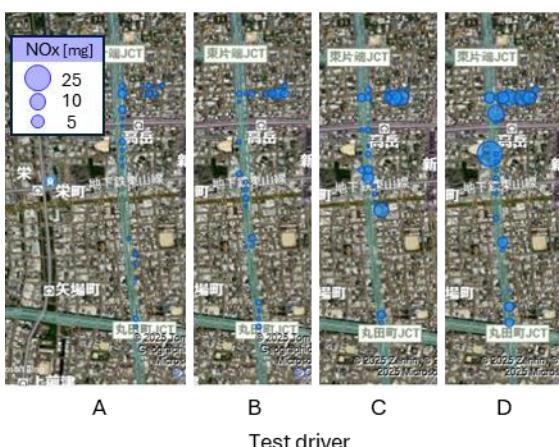


図3 ドライバごとの NOx 排出量・排出ポイント

### 2.2 ドライバの運転行動変容による低減効果

実路での排出ガス低減に関する先行研究[4]で実施した試験結果を図4に示す。この試験では、実路と同じ名古屋空港線を再現したデジタルツイン環境において、3名のドライバが複数回の走行試験を行った。特に、ドライバAは運転操作に特徴があり、前車を追従走行する際、アクセル操作をON/OFFして速度を調整する傾向が確認された。また、加速時にアクセルの踏み込み量が大きくなり、排出ガスが増加するような傾向も認められた。

この試験結果を基に、パワートレーン・排気システムの制御改良、ドライバの運転行動変容、交通流システムの最適化による排出ガス低減量の内訳を解析した結果を図5に示す。アクセル操作のON/OFFによるエンジンの間欠停止・再始動や、アクセル踏み込み量が大きい際のエンジン高負荷時におけるパワートレーン・排気システムの制御による効果が▲53%と最も大きかった。次いで、ドライバの運転特性による排出ガス差が大きく、NOx排出量の多いドライバの運転行動を、排出量の少ないドライバの運転操作に近づけることができれば、NOxを▲28%低減することができる。さらに、交通流の影響によるNOx低減効果は▲13%であり、クルマ・ドライバ・交通流を三位一体で最適化することで、実路走行時のNOx排出量を最大で▲94%低減することが可能である。

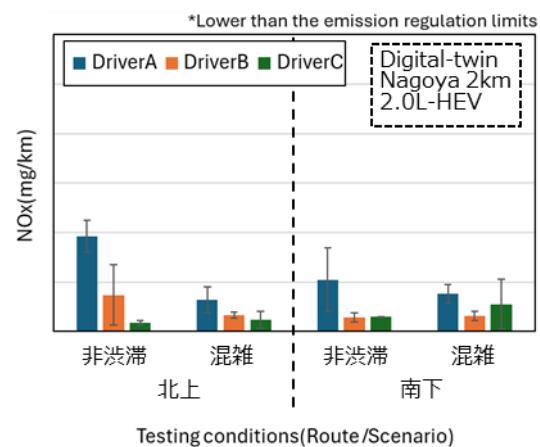


図4 ドライバ・交通流シナリオごとの NOx 排出量

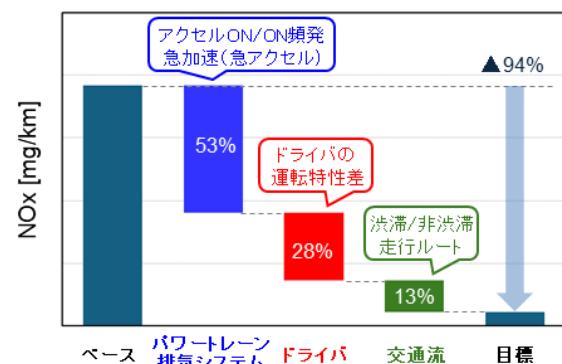


図5 実路における NOx 排出低減量の内訳

### 3. 試験装置・試験プロトコル

#### 3.1 デジタルツイン評価環境

ドライバの運転行動変容の働きかけによる排出ガス低減効果を評価するためには、ドライバや手段ごとの交通状況を揃える必要があるが、実路での走行では交通状況を制御することができないため、本研究では図 6 に示すような交通流シミュレーションとドライビングシミュレータを連携した、デジタルツイン評価環境を構築した[3]。

PTV 社製のマルチモーダル交通流シミュレーションソフト Vissim[5]を用いて、交通量やドライバの運転特性などを設定して任意の交通流シナリオを作成する。作成した交通流シナリオは Unreal Engine[6]で構築した 3D 都市モデル上に描画し、ドライビングシミュレータで走行環境を再現する。ドライビングシミュレータは 202 度のシムピットスクリーンシステム[7]とサウンドシステムで構成し、視覚・聴覚による没入感を提供する環境とした。ドライビングシミュレータの運転操作は XILS (X In the Loop Simulation) に入力され、実車と同様のエンジンおよびハイブリッド制御が動作する。車両シミュレーション CarMaker[8]にて車速や自車位置が計算され、その結果がドライビングシミュレータの描画や交通流シミュレーションに反映される。このサイクルを繰り返すことで、実路と同等の環境での走行試験が可能となる。さらに、各状態量から NOx 排出量を推定するエミッションモデルを用いて NOx の排出量を計算した。



図 6 デジタルツイン評価環境のシステム構成図

#### 3.2 ドライバへの行動変容促進手段

本研究では、表 2 に示す手段によってドライバの認知バイアスに働きかけすることで、運転行動の変容促進を図った。

##### A 介入なし

ドライバは自身の運転による NOx 排出を把握することなく、交通流に合わせて自由に運転する。

##### B NOx 表示アプリ (アンカリング効果)

NOx 排出量に応じてバーが動き、排出量が増加するとバーの色が赤に変化することでドライバに基準を提示し、NOx の排出を抑制する運転行動への変容を促す。

##### C ロボット (同乗者効果[9], 社会規範)

NOx 排出量が閾値を超えた際、ロボットがドライバに NOx 排出したことを通知する。ロボットによる発話により、NOx 排出を抑制する心理・行動への変容を促す。ロボットは先行研究[9]に倣い SHARP 製 RoBoHoN[10]を使用した。

##### D ロボット×生成 AI

###### (類似性バイアス, カクテルパーティー効果)

事前に入力した趣味嗜好を基に、図 7 に示す AWS アーキテクチャの生成 AI を活用して発話内容を作成した。発話内容をパーソナライズすることで、働きかけに対する受容性を高めるとともに、発話内容が変化することで、飽きることなく理想的な運転行動の定着を図る。

表 2 ドライバの運転行動変容手段

認知バイアス	期待する効果
 <b>NOx表示アプリ</b> 最初に表示された数値・情報が、その後の判断や意思決定に影響を与える	NOx排出に応じたバーの表示 排出量の増加で色が赤くなる 運転によるNOx排出を認識し、バーの大きさや色を基準にして運転行動を変容する
 <b>ロボット</b> <b>同乗者効果</b> 同乗者の存在が行動・心理に影響する現象 <b>社会規範</b> 暗黙の規則・期待が行動に影響する現象	NOx排出が閾値を超えるとロボットがNOx排出をお知らせ ロボット(同乗者)からの指摘で運転行動を変容する
 <b>ロボット×生成AI</b> <b>類似性バイアス</b> 共通点が心理的な親近感を高める <b>カクテルパーティー効果</b> 自分の興味のある話題など、特定の情報だけが耳に入りやすい心理的現象	生成AIで趣味嗜好に基づいた発話内容作成、NOx排出時にロボットが発話する 趣味嗜好に関する発話内容で働きかけに対する受容性を向上 理想の運転行動を定着させる

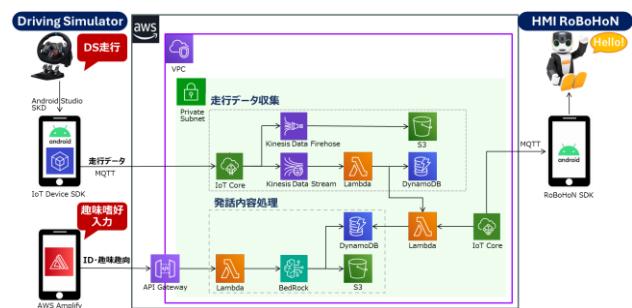


図 7 AWS システムアーキテクチャ

### 3.3 試験プロトコル[表 3]

本研究では、被験者として 20 代～50 代の 22 名を選定した。今回選定した被験者は筆者と同じ部署に所属し、通勤に自動車を使用し、排出ガスに関する基本的な知識を有している。各被験者に対し、A 介入なし、B NOx 表示アプリ、C ロボット、D ロボット×生成 AI を実施した。試験順序によるドライビングシミュレータでの運転操作慣れの影響を排除するため、A 介入なしを最初に実施し、その後 B、C、D の試験順序はランダムに設定した。

走行試験は、デジタルツイン環境の名古屋空港線で実施した。交通流シミュレーションにより渋滞のないスムーズな交通流シナリオを作成し、ドライバごとに交通状況の差が出ないように、走行試験前に交通流シミュレーションおよび Unreal Engine を再起動してから試験を開始した。

試験データとしてドライビングシミュレータで計測したアクセル開度などの運転操作に加え、エンジン ECU から取得した状態量を基にエミッションモデルで推定した NOx 排出量を取得し、試験後には行動変容への働きかけに対する主観評価をアンケートにて取得した。

なお、本研究におけるドライバ特性が影響する評価については、トヨタ自動車の研究倫理審査委員会の承認を得て実施した。

表 3 試験プロトコル

被験者	22人 (20代～50代)
試験日時	'25/11/13～'25/11/19
促進手段 試験順序	A : 介入なし B : NOx表示 C : ロボット D : ロボット×生成AI #1 #2～#4 ランダム
試験方法	ドライビングシミュレータ 名古屋空港線 約2km (往復) 渋滞無し
評価項目	NOx排出量 [mg/km] 車速・アクセル開度、など 主観評価アンケート (試験後)

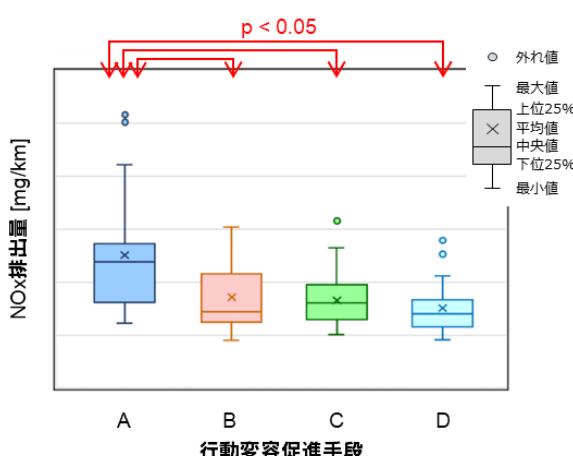


図 8 NOx 排出量に対する行動変容促進の効果

### 4. 試験結果と考察

#### 4.1 ドライバの行動変容促進による排出ガス低減結果

図 8 に、各行動変容促進手段を実施した際の NOx 排出量の比較を示す。A 介入なしに対して、B NOx 表示アプリ、C ロボット、D ロボット×生成 AI と狙い通り NOx 排出量が低減する傾向が確認された。今回の結果について、行動変容促進手段による差が統計的に有意であるかを Friedman 検定にて確認した。その結果、A 介入なしに対して、B・C・D の各促進手段は有意に NOx 排出量が低減していたが、B・C・D の促進手段の間には有意差は認められなかった。

今回の試験の中で、顕著に効果がみられた被験者の各促進手段での運転における運転操作として図 9 に時系列のアクセル操作、図 10 に平均アクセル開度と平均車速の比較を示す。A 介入なしの運転では、NOx が排出するような急加速が多く見られたが、行動変容の働きかけがされた B・C・D での運転では、急峻なアクセル操作の回数が減少し、平均的にアクセル開度が低くなっていることから、狙い通り、働きかけによって NOx の排出を抑制するような運転行動に変容したと考えられる。

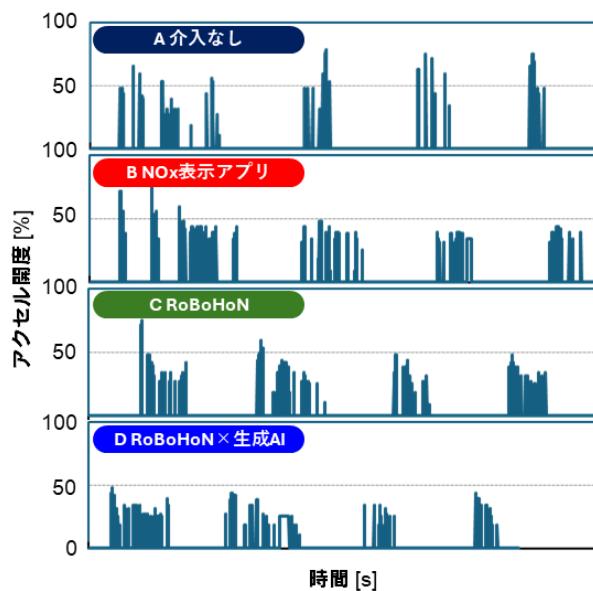


図 9 運転行動の促進手段によるアクセル操作の比較

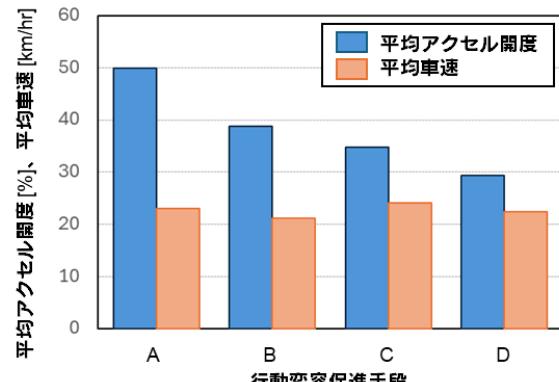


図 10 働きかけに対する運転操作の変化

## 4.2 ドライバの主観評価

試験後に実施したアンケートの結果を表 4 および図 11 に示す。各被験者のコメントからも、行動変容の促進によって「NOx 排出を抑えるように意識して運転する」など、運転行動が変容したことが確認できる。

B NOx 表示アプリでは、NOx 排出量が多くなるとバーが大きくなり赤く表示されることで、ドライバが“基準=アンカー”を持つことができたと考えられる。一方、「バーの表示に集中して運転に集中ができない」や「画面が見にくかった」など、バーの表示だけでなく音や振動と組合せるなど、ゼロエミッションを目指す上でさらなる検討が必要であることがわかった。

C ロボットの発話に対しては、「NOx が排出しないように意識して運転した」というポジティブな回答と、「発話が運転の妨げになる」といった改善を要する意見の両方が見られた。先行研究[9]では、日常生活でロボットを愛用する被験者に対して“同乗者効果”が確認されたが、今回は初めてロボットを見る被験者も多かったため、“同乗者効果”や“社会規範”的効果は得られにくかったと考えられる。HMI (Human Machine Interface) や UI (User Interface) をパーソナライズ化するなどの工夫が必要である。

D ロボット×生成 AI では、C の単純な NOx 排出の指摘に比べ、興味を示す被験者は多かった。しかし、ドライバの認知・判断への働きかけが狙いだったのに対して、「発話内容の理解に時間がかかった」など、認知に時間がかかってしまったドライバについては、判断・操作への効果は得られなかつたと考えられる。また、ドライビングシミュレータでの排出ガス試験という特殊な環境では、日常運転時に効果を狙った“類似性バイアス”や“カクテルパーティー効果”は得られにくかったと考えられる。AI エージェントを活用した対話も検討する必要がある。

表 4 試験後のアンケート結果

	行動変容促進効果	改善ポイント
	NOxのバーが赤くならないよう気を付けて走れた 運転操作に対してバーが変わるのでわかりやすかった  アンカリング効果	NOxバーが気になって運転に集中できなかった スマホの画面が見にくかった 基準線があった方が分かりやすい
	ロボットから注意されないよう運転に気を付けられた NOx排出しないように運転することができた  同乗者効果・社会規範	普通に走行した時に注意されて躊躇しかった 毎回同じ言葉で飽きそう 急に喋りだしてびっくりした バー表示もあった方が分かりやす
	次に何を喋るのか気になった 発言内容が変わるので注意が引かれた 発話によりポジティブな気分で運転することができた  類似性バイアス カクテルパーティー効果	発話内容を理解するのに時間がかかった 何を言ってるかわからなかった 注意コメントだけでなく褒めたり雑談があるといい

## 4.3 行動変容促進手段に対する考察

今回の試験では、通勤などで日常的に運転を行い、自動車の排出ガスに関する基本的な知識を有する被験者を対象とした結果である。図 12 に示す行動変容ステージにおいて、準備期・実行期・維持期に属するグループであったため、B・C・D の促進手段によらず、NOx 排出量を「見える化」することによって十分な効果が得られたと推察される。

一方で、本プロジェクトで対象とするドライバの大半は、自動車からの排出ガスに関する知識は少なく、行動変容ステージの無関心期・関心期に属するものと考えらえる。そのため、今回のような NOx の排出をドライバに知らせるような行動に対する働きかけを実施する前に、自動車からの排出ガスや環境に関して、考え方に対する働きかけを実施するか、ゲーム要素を持たせて NOx 排出を間接的に表現し、ドライバが直感的に運転操作を変容させるような工夫も必要になると考えられる。

さらに、今回のような単発的な試験ではなく、NOx 排出を抑制する運転行動の長期的な定着も大きな課題である。今回の試験では、D ロボット×生成 AI による NOx 低減効果に有意差は見られなかったが、C ロボットの発話では慣れによって効果が低下する可能性があるため、生成 AI や AI エージェントを活用した発話のパーソナライズ化は有効な手段であると考えられる。

今後は、発話内容だけでなく、発話タイミングなどドライバの注意を引く手法の検討や、より長期間にわたる効果検証が課題となる。

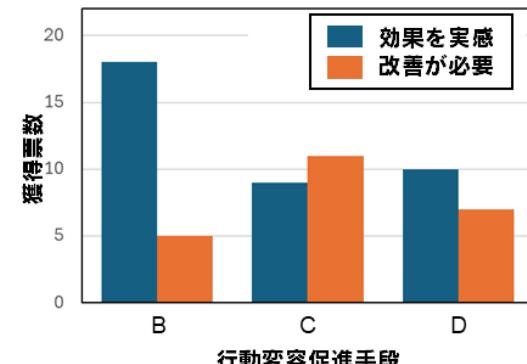


図 11 主観アンケート結果

行動変容ステージ	環境に対して知識・興味がない	自分には関係ないと思っている(傍観者)	行動したいが何をすればいいかわからない	できる範囲で行動している	積極的に情報を得て行動している 目標指すステージ
	無関心期	関心期	準備期	実行期	維持期
行動の阻害要因(ボトルネック)	・環境への影響の認知が不足している ・環境に対する情報発信が不足している	・自分がだけエミッション排出しているかわからっていない ・メークが取引相手の、自分は貢献できる事はない、という認識(傍観者)	どうすれば環境にいい影響を与えるかわからず 一歩づかがわればやる	環境に優しい運転をしているつもりでも、どうくい貢献できていないかわからず	
戦略・方針	情報発信を強化 エミッション排出ポイント 運転結果の振り返り	エミッション排出量視える化 (燃費と同) 運転で貢献できることをデータで示す (理想状態・他者と比較)	運転アドバイス (事前学習・振り返り) 簡単にできるツールを提供 生成AIで指摘・警める	一緒に活動 環境にいい運転を模範化 →好事例として紹介	
認知バイアス	情報提供(権威バイアス) 社会比較(同調性) 単純接觸効果	アンガリング・ティフルト 情報提供(権威バイアス) 社会比較(同調性)	アンガリング・ティフルト 社会比較・社会規範 インセンティブ	社会規範 イケア効果 コモトメント	
考え方に対して働きかけ			行動に対して働きかけ		

図 12 自動車からの排出ガスに対する行動変容ステージ

## 5. まとめ

本研究では、リアルワールドゼロエミッションの実現を目指し、ドライバの運転行動変容を促進する介入手段の効果をデジタルツイン環境を用いたドライビングシミュレーション試験により評価した。

- NOx 排出量の「見える化」やロボット・生成 AI による対話的介入など、複数の行動変容促進手段を導入し、運転行動の変容による排出ガス低減効果を検証した。その結果、いずれの促進手段においても、介入なしの場合と比較して NOx 排出量の有意な低減傾向が確認され、NOx 排出量の「見える化」によって、ドライバの運転行動が排出ガス排出抑制方向へ変容することが示された。
- 試験後のアンケート調査からも、被験者が排出ガスを意識した運転行動へと変化したことが確認された。一方で、促進手段の種類による効果の差は限定的であり、ドライバの知識や関心度に応じたアプローチや、長期的な行動定着を目指した工夫が必要であることが示唆された。
- 今後は、より幅広いドライバ層への適用や、発話内容・タイミングの最適化、長期間にわたる効果検証などを進めることで、持続的な排出ガス低減を着実に進めていく。

## 謝 辞

本研究の実施にあたっては、トヨタテクニカルディベロップメント 森部様、PTV ジャパン 三浦様をはじめとする多くの方々から多大なるご支援とご協力を賜りました。システム構築および評価に貢献いただき、心より感謝を申し上げます。

## 参考文献

- [1] 高須賀 蓮：日本におけるカーボンニュートラル燃料を使用した自動車のライフサイクル GHG 排出量予測、第 35 回内燃機関シンポジウム、予稿集、NO24-51 (2024)
- [2] “Regulation - 2024/1257 - EN - EUR-Lex”，  
<https://eur-lex.europa.eu/eli/reg/2024/1257/oj>
- [3] 池本 雅里：リアルワールドエミッション低減技術開発（第 1 報）—デジタルツイン開発環境の構築—、自動車技術会 2025 年秋季大会、学術講演会公演予稿集、文献番号 20256253、NO.253 (2025)
- [4] 峯尾 健太郎：リアルワールドエミッション低減技術開発（第 2 報）—デジタルツイン開発環境を活用したエミッション低減検討—、自動車技術会 2025 年秋季大会、学術講演会公演予稿集、文献番号 20256254、NO.254 (2025)
- [5] “交通シミュレーションソフトウェア | PTV Vissim”，  
<https://www.ptvgroup.com/ja/products/ptv-vissim>
- [6] “Unreal Engine | Epic Games”，  
<https://www.unrealengine.com/ja>
- [7] “シムピットスクリーンシステム | エイムス株式会社”，  
<https://www.aims-jp.net/>
- [8] “CarMaker | IPG Automotive 株式会社”，  
<https://www.ipg-automotive.com/jp/products-solutions/software/carmaker/>
- [9] “運転行動変容を促すドライバエージェント研究”，  
<https://www.potstill.net/>
- [10] “モバイル型ロボット『RoBoHoN(ロボホン)』 | SHARP”，  
<https://corporate.jp.sharp/news/190218-d.html>

# AI とのコミュニケーションによる環境配慮行動の誘導可能性 —ごみ処理ゲームを題材とした検証—

日室聰仁<sup>†1</sup> 後藤晶<sup>†2</sup> 安藤香織<sup>†3</sup>NEC ソリューションイノベータ株式会社<sup>†1</sup> 明治大学<sup>†2</sup> 奈良女子大学<sup>†3</sup>

## 1. はじめに

気候変動対策の重要性は、地球温暖化を抑制し、自然環境や人間社会の持続可能性を確保するために不可欠である。気候変動を抑制するには、個人の環境配慮行動が重要であり、個人の環境配慮行動をいかに促進するかが鍵となる。環境配慮行動を促進する一方法として、コミュニケーションが挙げられる。先行研究では人間同士で環境問題に関するコミュニケーションをとることによって環境配慮行動が促進されることが示唆されている<sup>[1][2]</sup>。コミュニケーションが効果を持つ理由としては、協力への期待が生まれる<sup>[3]</sup>、情報・規範の共有で全体の便益に目が向く<sup>[4]</sup>、集団アイデンティティが高まる<sup>[5]</sup>、などが先行研究によって示唆されている。

コミュニケーションは環境配慮行動を促進する有益な方法であるが、一方、この知見を社会実装する際には、複数の人間を同一時間に集めてコミュニケーションをとらせる必要があるため、実施コストが高いという問題がある。この問題を解決する方法として、筆者らは生成 AI に着目した。人間同士のコミュニケーションではなく人間と生成 AI とのコミュニケーションによって人間の環境配慮行動を促進されるのであれば、スマートフォンなどを活用して、いつでもどこでもコミュニケーションが可能となり、実施コストの問題を解決できると考えた。そこで、本活動では、AI とのコミュニケーションが環境配慮行動を促進するかどうかについて、ごみ処理ゲームを題材に検証した。また、AI にキャラクターを設定することで、人間の行動を誘導できる可能性があるかについて、生成 AI を活用したシミュレーションを通して検証した。

## 2. 方法

繰り返し公共財ゲームをごみ分別の文脈にアレンジしたごみ処理ゲームを活用して AI とのコミュニケーションが環境配慮行動を促進するかどうかを検証した。

繰り返し公共財ゲームとは、複数の参加者が何回かにわたりて公共財への貢献を選択する社会的ジレンマの実験・モデルである。各プレイヤーは毎回、与えられた資源の一部または全部を公共財（共有プール）に投入するか、自分の手元に残すかを決定する。公共財への合計貢献額は一定

の倍率をかけてプレイヤー全員に均等に分配される。各プレイヤーは毎回、与えられた資源の一部または全部を公共財（共有プール）に投入するか、自分の手元に残すかを決定する。公共財への合計貢献額は一定の倍率をかけてプレイヤー全員に均等に分配される。ゲームは一回きりではなく、同じメンバーで繰り返しラウンドを行う。各ラウンドの選択は次のラウンドに影響を与える可能性があるため、協力行動や裏切り行動、戦略的対応が発生する。

この繰り返し公共財ゲームの枠組みを環境問題の文脈のアレンジしたゲームがごみ処理ゲームである。各プレイヤーは工場の経営者としてゲームに参加する。工場では毎ラウンド 20kg のごみが発生し、ごみを資源ごみとして処理するか、可燃ごみとして処理するかの割合を各プレイヤーは選択する。資源ごみの処理には 1kgあたり 1 ポイントの費用を支払う必要があるが、集まった資源ごみの総量を 2 倍した数値がポイントとして次のラウンドにプレイヤー全員に均等に分配される。一方、可燃ごみはポイント不要で処理できる。本研究では 3 人で 1 グループを組みゲームを実施した。また、10 ラウンド繰り返しとした。ラウンド数については実験参加者には明かさずに実験を進めた。

### 2.1 実験 1 : AI とのコミュニケーション実験

実験 1 では、人間の実験参加者と生成 AI とでグループを組みごみ処理ゲームをプレイさせた。グループの構成は人間 1 人/AI2 人、人間 2 人/AI1 人、人間 3 人/AI0 人の 3 条件とした。また、5 ラウンド終了後にグループメンバーとチャットでコミュニケーションをとる条件とコミュニケーションをとらない条件の 2 条件を設けた。資源ごみ量を選択する画面を図 1、グループメンバーとコミュニケーションをとる画面を図 2 に示す。コミュニケーションはチャット形式で各プレイヤーは 3 回の発言機会が提供された。発言順序はランダムに決めるようにした。

生成 AI は OpenAI 社の GPT-4o を利用し、与えたプロンプトは図 3 とした。プロンプトは人間の実験参加者に与えている情報と同等の情報を与えるように設計した。ただし、1 ラウンド目の資源ごみの量のみ生成 AI の回答ではなく、過去に筆者らが実施したごみ処理ゲームにおける 1 ラウンド目の人の行動データからランダムに回答するようにした。この対応は人間の行動と生成 AI の行動を似せるための対応である。この対応を入れない場合、AI は常に資源ごみとして処理する行動をとることが確認されたため、本対応を実施した。この対応によりかなり人間の行動に近づくことが

Communication with AI guide pro-environmental behavior?

<sup>†1</sup> AKIHITO HIMURO, NEC Solution Innovators, Ltd.

<sup>†2</sup> AKIRA GOTO, Meiji University

<sup>†3</sup> KAORI ANDO, Nara Women's University

確認されている。

実験は oTree<sup>[6]</sup>を用いて行い、人間の行動の変化を確認することによってAIが人間の行動を誘導できる可能性があるかについて検証した。

1年目 ごみの処理方法を決める



図 1：資源ごみ量を選択する画面

グループメンバーとのチャット



図 2:コミュニケーション画面

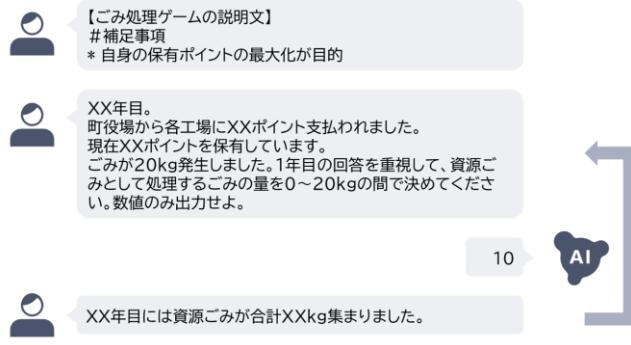


図 3：プロンプト

## 2.2 実験 2

実験 2 では、キャラクターを設定していない生成 AI1 人とキャラクターを設定している生成 AI2 人でごみ処理ゲームをプレイし、5 ラウンド終了後に AI 同士でコミュニケーションをとり、その後、さらに 5 ラウンドを続けた。コミュニケーション時は 3 回の発言機会が提供され、発言順序はランダムに決定された。

生成 AI は OpenAI 社の GPT-4o を利用し、与えたプロンプトおよびラウンド 1 の挙動は実験 1 と同様とした。キャラクターを AI に設定するタイミングはコミュニケーション開始直前とし、「XXX になり切って行動せよ」と AI に指示を与えた。キャラクター設定は利己的と利他的の 2 種類とし、XXX に「利己的」、「利他的」を設定するようにした。CTRL 条件としてキャラクター未設定のコミュニケーションあり・なし条件のデータも収集した。

キャラクターを設定していない AI の行動を観察することによって、キャラクターを設定した AI が行動を誘導できる可能性があるかについて検証した。

## 3. 実験結果

### 3.1 実験 1：AI とのコミュニケーション実験

2025 年 1 月 24 日に Yahoo クラウドソーシングを活用して実験参加者を集めて実験をした。実験参加者はランダムに条件に振り分けされた。人間実験参加者の有効データ数は 452 人/230 グループ(男性 308 名、女性 139 名、無回答 5 名、年齢 M=50.48, SD=11.41, 年齢未回答者を除く)であり、各条件の内訳は図 4 の通りである。

	コミュニケーション あり条件	コミュニケーション なし条件
人間1人/AI2人 条件	40人/40グループ	42人/42グループ
人間2人/AI1人 条件	76人/38グループ	72人/36グループ
人間3人/AI0人 条件	114人/38グループ	108人/36グループ

図 4：実験 1 の人間実験参加者の有効データ数内訳

各条件の各人の資源ごみ量平均の推移を図 5、コミュニケーション機会が提供される前のラウンド(1~5 ラウンド：前半ラウンド)とコミュニケーション機会が提供された後のラウンド(6~10 ラウンド：後半ラウンド)の人間の資源ごみ量平均の差分を図 6 に示す。

AI とのコミュニケーションによって資源ごみ量が増加することが確認された。また、人間だけのグループに比べて、AI が含まれるグループのほうが後半の資源ごみ量が多いこと、AI の人数に比例して増加する傾向があることが確認された。

また、資源ごみ量を目的変数として、各条件および前半後半を説明変数として、人間 3 人 AI0 人 × 前半ラウンド × コミュニケーションなしを切片として、一般線形混合モデルでの分析した結果を表 1 に示す。コミュニケーションあり × 後半ラウンドで 0.81 ポイント資源ごみ量を有意に増やしていること、さらに AI1 人グループの場合 0.94 ポイント、AI2 人グループの場合 1.87 ポイント、資源ごみ量を有意に増やしていることが確認された。

これらの結果より、AI とのコミュニケーションによって環境配慮行動を促進する可能性があるが示唆された。また、人間同士のコミュニケーションより AI とのコミュニケーション

ヨンのほうが環境配慮行動を促進する可能性があること、AI の人数に比例して環境配慮行動を促進される可能性があることが確認された。

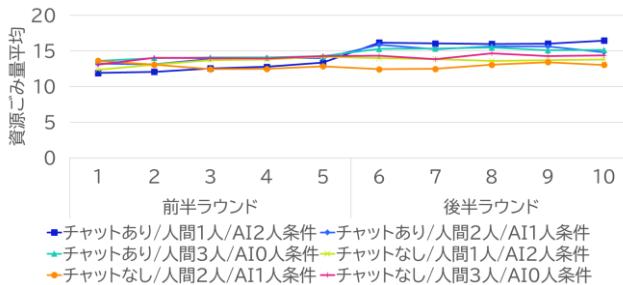


図 5：実験 1 条件別の資源ごみ量平均の推移

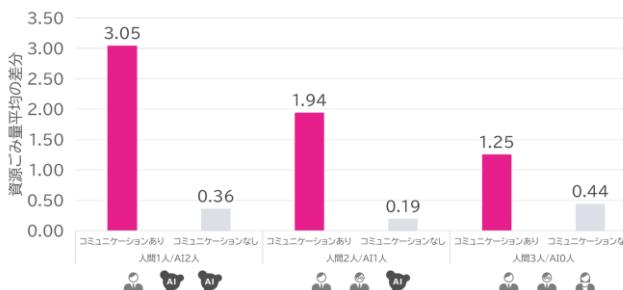


図 6：実験 1 前半と後半の資源ごみの量平均の差分

表 1：実験 1 一般線形混合モデル分析の結果

Predictors	リサイクル量	
	Estimates	p
<b>切片</b>	<b>13.85 (12.51 – 15.18)</b>	<b>&lt;0.001</b>
人間2人/AI1人	-1.43 (-3.43 – 0.58)	0.163
人間1人/AI2人	0.39 (-1.81 – 2.59)	0.728
コミュニケーションあり	0.16 (-1.71 – 2.02)	0.87
<b>後半ラウンド</b>	<b>(-0.01 – 0.90)</b>	<b>0.056</b>
人間2人/AI1人 ×	1.24	0.383
コミュニケーションあり	(-1.55 – 4.04)	
人間1人/AI2人 ×	-2.29	0.151
コミュニケーションあり	(-5.41 – 0.83)	
人間2人/AI1人 ×	-0.25	0.498
後半ラウンド	(-0.97 – 0.47)	
人間1人/AI2人 ×	-0.08	0.854
後半ラウンド	(-0.94 – 0.78)	
コミュニケーションあり ×	<b>0.81</b>	<b>0.012</b>
後半ラウンド	<b>(0.18 – 1.44)</b>	
人間2人/AI1人 ×	<b>0.94</b>	<b>0.067</b>
コミュニケーションあり × 後半ラウンド	<b>(-0.06 – 1.94)</b>	
人間1人/AI2人 ×	<b>1.87</b>	<b>0.003</b>
コミュニケーションあり × 後半ラウンド	<b>(0.65 – 3.09)</b>	

### 3.2 実験 2：AI による誘導可能性実験

OpenAI 社の GPT-4o を利用して実験した。データ数は各条件 50 グループのデータをサンプリングした。各条件のキャラクターを設定していない AI の資源ごみ量平均の推移を図 7、コミュニケーション機会が提供される前のラウンド(1~5 ラウンド: 前半ラウンド)とコミュニケーション機会が提供された後のラウンド(6~10 ラウンド: 後半ラウンド)の

キャラクターを設定していない AI の資源ごみ量平均の差分を図 8 に示す。

また、資源ごみ量を目的変数として、各条件および前半後半を説明変数として、CTRL チャットなし × 前半ラウンドを切片として、一般線形混合モデルでの分析した結果を表 2 に示す。CTRL チャットあり × 後半ラウンドの場合 3.29 ポイント、利他的条件 × 後半ラウンドの場合 2.88 ポイント、資源ごみ量を有意に増やすことが確認された。また、利己的条件 × 後半ラウンドの場合 0.9 ポイント資源ごみ量を有意に減らすことが確認された。

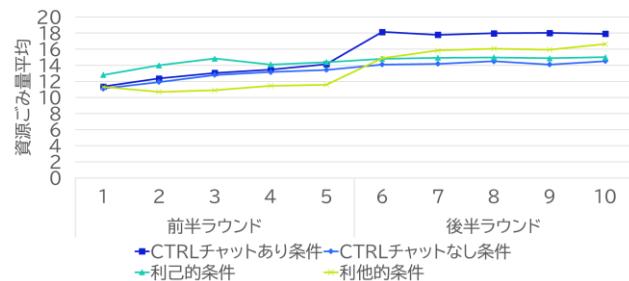


図 7：実験 2 条件別の資源ごみ量平均の推移



図 8：実験 2 前半と後半の資源ごみの量平均の差分

表 2：実験 2 一般線形混合モデル分析の結果

Predictors	リサイクル量	
	Estimates	p
<b>切片</b>	<b>12.48 (11.00 – 13.97)</b>	<b>&lt;0.001</b>
CTRLチャットあり条件	0.4 (-1.70 – 2.50)	0.708
利他的条件	-1.29 (-3.39 – 0.81)	0.229
利己的条件	1.54 (-0.55 – 3.64)	0.149
<b>後半ラウンド</b>	<b>1.8 (1.14 – 2.46)</b>	<b>&lt;0.001</b>
CTRLチャットあり条件 ×	<b>3.29 (2.36 – 4.22)</b>	<b>&lt;0.001</b>
後半ラウンド	<b>2.88 (1.95 – 3.82)</b>	<b>&lt;0.001</b>
利他的条件 ×	<b>-0.9 (-1.84 – 0.03)</b>	<b>0.057</b>
後半ラウンド		

これらの結果より、利他的なキャラクターを設定された AI とグループを組みコミュニケーションをとる場合は環境配慮行動が促進される可能性があること、利己的なキャラクターを設定された AI とグループを組みコミュニケーションをとる場合は環境配慮行動が抑制される可能性があること、が示唆された。また、キャラクターを設定していない CTRL 条件チャットあり条件が後半ラウンドに 3.29 ポイント資源ごみ量を有意に増加させていることより、キャラク

ターを設定していない OpenAI 社の GPT-4o は利他的な側面をデフォルトで有している可能性があると考えられる。

#### 4. 考察

本研究では、生成 AI とのコミュニケーションが人間の環境配慮行動を促進し得るか、また AI へのキャラクター設定が行動誘導にどのように作用するかを、ごみ処理ゲームを用いた実験を通じて検証した。

実験 1 の結果から、AI とコミュニケーションをとることで、資源ごみ量が有意に増加することが認められ、AI の人数に比例して促進効果が高まることが示唆された。これは、従来の人間同士のコミュニケーションによる協力促進効果だけでなく、AI が参加者に対して「利他性」の規範的モデルを提示し、環境配慮行動への期待や集団アイデンティティを形成し得る可能性を示している。AI が単なる情報提供者として機能するのではなく、社会的存在として、コミュニケーションを通じて人間の社会的ジレンマに影響を与えることは新たな知見と考える。

さらに、実験 2 の結果では、AI にキャラクター（利他的・利己的）を設定することで、「利他的」なキャラクターが資源ごみ量を増加させ、「利己的」なキャラクターが資源ごみ量を減少させるといった、行動誘導効果が明確に観察された。これは AI の振る舞いが、参加者の意思決定に意味のある影響を与えていていることを示し、AI 設計においてキャラクター性が行動誘導の鍵となる可能性が高いことを示している。また、キャラクター未設定時においても利他的な側面がデフォルトで現れる傾向があったことは、生成 AI のアルゴリズムや設計思想が人間の協力行動を自然と引き出す性質を持つのか、この事例では偶然利他的に AI が行動したことに起因するのか、さらなる分析を要するポイントと考える。

社会実装面から考えると、AI とのコミュニケーションが人間同士のコミュニケーションに比べて安価かつ容易に大規模展開できる利点があり、持続可能な社会を目指す環境教育やまちづくり、廃棄物分別促進など様々な分野で有効活用が期待される。加えて、AI の性格付与によりターゲット層や目的に合わせた個別のコミュニケーション設計が可能となるため、今後の AI 応用設計の指針となる知見が得られたといえる。

一方で、本研究の限界として、ごみ処理ゲームという抽象化されたモデルを用いている点、AI のキャラクター設計やプロンプト設定が現実の多様な状況に完全には一致しない可能性、実験参加者の属性やゲーム内行動が現実社会での廃棄物分別行動と直結するとは限らない点が挙げられる。また、実験 2 では人間を対象とした実験ができていない点も課題と考える。今後は、現実の環境配慮行動と AI コミュニケーションの効果検証、および AI の発言内容・コミュニケーションの質的分析など、さらなる研究が必要である。

#### 5. おわりに

本研究は、生成 AI とのコミュニケーションが人間の環境配慮行動を促進し得ること、その効果は AI の人数やキャラクター設定により変化し得ることを、ごみ処理ゲームを用いた実験を通じて示した。特に、AI による利他的な行動モデルの提示とキャラクター設定は、協力促進や行動変容につながる有効な手段であることが示唆された。

これらの知見は、現実社会において AI を活用した環境啓発やまちづくり支援、行動変容プログラムの開発につながる基礎的情報を提供するものである。

今後は、より実践的な場面での AI コミュニケーションの効果検証や、多様な属性を持つ AI の設計と社会応用、公平性・倫理性の観点からの課題整理など、多方面での研究の深化が望まれる。

#### 参考文献

- [1] 安藤香織, 杉浦淳吉, 大沼進, 安達菜穂子. “環境配慮型ライフスタイルを促進するコミュニケーションの検討：国際比較実験を用いて”, 環境社会心理学研究, 24巻, 1-100.
- [2] 早瀬光司.“実団地における資源循環型ライフスタイル普及のための環境コミュニケーションとその効果に関する実証的研究”, 環境省廃棄物処理技術情報, K1604.
- [3] 大沼進, “「廃棄物処理ゲームと二次的ジレンマ」広瀬幸雄(編)『シミュレーション世界の社会心理学』, pp. 149-152, ナカニシヤ出版.
- [4] Dawes, R. M., Alphons, J. C., Kragt, A. J., & Orbell, J. M. (1990). Cooperation for the benefit of us: Not me, or my conscience. In J. J. Mansbridge (Ed.). *Beyond self-interest* (pp. 97-110). The University of Chicago Press
- [5] Kerr, N. L., & Kaufman-Gilliland, C. M. (1994). Communication, commitment, and cooperation in social dilemma. *Journal of Personality and Social Psychology*, 66(3), 513. <https://doi.org/10.1037/0022-3514.66.3.513>
- [6] Chen, D. L., Schonger, M., & Wickens, C. (2016) 「oTree—An open-source platform for laboratory, online, and field experiments」 *Journal of Behavioral and Experimental Finance*, 9, 88-97. <https://doi.org/10.1016/j.jbef.2015.12.001>

#### 脚注

本稿の実験 1 について、2025 年 8 月 27 日に開催された第 12 回 気候変動・省エネルギー行動会議 BECC JAPAN 2025、および、2025 年 9 月 20-21 日に開催された日本社会心理学会第 66 回大会 WS08 ワークショップで口頭発表済みである。

# 運動予測可能性への選好に基づく ASD リスク発見の可能性

大森幹真<sup>†1</sup>・大塚千冬<sup>†2</sup>・國本彩永<sup>†2</sup>・伊藤愛華<sup>†2</sup>・木内愛奈<sup>†2</sup>・東美穂<sup>†1</sup>  
早稲田大学人間科学学術院<sup>†1</sup>・早稲田大学人間科学部<sup>†2</sup>

## 1. はじめに

自閉スペクトラム症(ASD)の早期発見方法として視線機能計測が使用されている(e.g., Omori [1], Pierce et al., 2011)。Omori[1]は非言語的課題を用いて ASD リスク児は予測可能な動き(運動予測可能性)をする图形刺激に対しての視覚的選好が、徐々に増加することを示し、実践現場での応用が期待されている。一方で、運動予測可能性を用いて未診断者を分類し、他の視線機能が先行研究と同様の視線パターンを示すかは未解明である。そのため、本研究では先行研究をもとに判別関数を作成し、それをもとに未知データとして取得した 56 名の参加者を ASD リスク群と非リスク群に分類するときの精度を評価することを目的とした。さらには、分類後の視線パターンの分析を行い、先行研究と同様の傾向が見られるかを明らかにすることも目的とした。

## 2. 方法

### 2.1 参加者

未診断の子どもから大人合計 56 名が新規に参加した。本研究は早稲田大学倫理審査委員会(2024-607)での承認後、参加者本人に説明・書面同意を得た上で、研究を実施した。

### 2.2 刺激と装置

運動予測可能性を評価するために、图形を一筆書きでなぞる P 刺激と、ランダムな順序でなぞる U 刺激を用意した。刺激提示方法では P 刺激と U 刺激を画面上の左右いずれかに同時に提示する選好注視法を使用した。1 試行の提示時間は 10 秒であり、各参加者に 6 試行ずつ実施した。刺激の提示用にノートパソコン(PC)と視線機能を計測するためのアイトラッカー(Pro Fusion, Tobii technology Japan)を使用した。

### 2.3 手続き

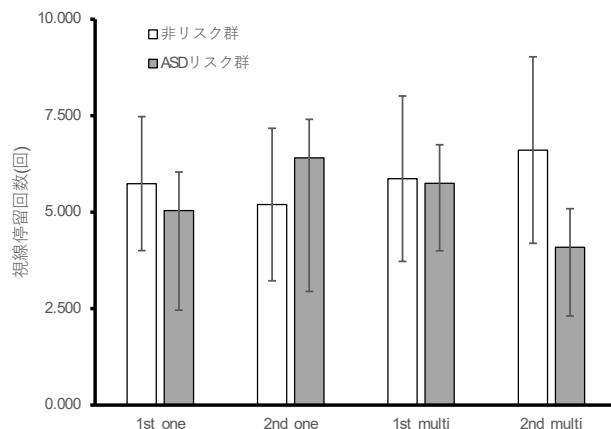
- (1) 判別関数の算出: Omori[1]から、運動予測可能性への選好による ASD リスク判別する関数を SPSS により算出した。
- (2) 視線機能計測: 参加者はアイトラッカーの画面から 60cm の位置に着席し、キャリブレーションを行った。選好注視法により画面上に 2 つの動画が同時に提示され、合計 6 ペアの動画を参加者には自由に観察することを求めた。
- (3) 従属変数: 10 秒間の提示を前後半 5 秒ずつに分けた。そのうえで刺激に対しての①総注視時間、②視線停留回数を用意した。
- (4) 訓練データの判別分析および、未知データの予測をし、

その後に未知データ内での各変数の分散分析を行った。

## 3. 結果

判別関数は刺激提示の前後半での P 刺激への選好割合の変化 $[F(1,26)=7.59, p=.011, \text{Wilks's } \lambda = .774, \text{ 固有値} = .29, \text{ 正準相関} = .48]$ で分類可能であり、群間の差異を説明する上で中程度の関連性があった。これを元に未知データの判別分析を行ったところ、71.9%の精度で 56 名中 11 名(21.2%)に ASD リスクがある可能性を示した。その分類を元に分散分析を行ったところ、先行研究[1]と同様の視線パターン示した。特に視線停留回数において有意な交互作用が見られた $[F(1,50)=35.45, p=.000, \text{partial } \eta^2 = .42]$ 。下位検定を行ったところ、ASD リスク群において後半での U 刺激への注視回数が非リスク群に比べて低減し $[F(1,50)=10.02, p=.003, \text{partial } \eta^2 = .17]$ 、群内での P 刺激への注視回数も増加することを示した $[F(1,50)=11.32, p=.001, \text{partial } \eta^2 = .19]$ 。

図 1. 両群における刺激ごとの視線停留回数の推移



## 4. 考察

本研究の結果から運動予測可能性の視覚的選好の推移から ASD リスク児を予測できる可能性を示した。一方で、他の予測モデルを作成することによる予測精度や障害傾向との一致率をさらに向上させていく必要がある。

## 参考文献

- [1] Omori, M. (2025). Increased observation of predictable visual stimuli in children with potential autism spectrum disorder. *Scientific reports*, 15(1), 4572.
- 本研究は JSPS 科研費 (No. 22K13739) の助成を受けて行った。本研究に関して、開示すべき利益相反関連事項は存在しない。

†1 MIKIMASA OMORI, Faculty of Human Sciences, Waseda University

†2 CHIFUYU OTSUKA, Faculty of Human Sciences, Waseda University

†3 SAE KUNIMOTO, Faculty of Human Sciences, Waseda University

†2 AIKA ITO, Faculty of Human Sciences, Waseda University

†2 AINA KIUCHI, Faculty of Human Sciences, Waseda University

†1 MIHO HIGASHI, Faculty of Human Sciences, Waseda University

# 一般運転者の避讓行動促進に向けた VR 緊急走行体験システムの開発

西浦裕生<sup>†1</sup> 山口晋吾<sup>†2</sup> 田中章夫<sup>†2</sup> 山添大丈<sup>†1</sup>  
兵庫県立大学<sup>†1</sup> 姫路市消防局<sup>†2</sup>

## 1. はじめに

本稿では、一般運転者が緊急車両の運転を VR で体験することを通じて、緊急走行時に道を譲る「避讓行動」への理解と協力意識を高めることを目的として開発した VR 緊急走行体験システムについて述べる。

救急車や消防車などの緊急車両の緊急走行は、傷病者の搬送や火災現場への迅速な到着に不可欠であるが、一般車両によって妨げられる事例も少なくない。小池らによる名古屋市消防局の調査では、走行阻害要因として「不適切な避讓行動」が上位に挙げられている[1]。

一方、近年 VR による他者視点の体験が、共感や利他性、向社会的行動を高める効果を持つことが多分野で示されており[2,3]、VR は社会的行動を形成するメディアとして注目されている。これらの知見は、緊急車両の運転者視点を一般運転者に疑似的に体験させることで、避讓行動への理解と協力意識を高める教育的アプローチが有効である可能性を示唆している。

我々はこれまで、消防士のための緊急走行安全教育システムの実現を目指し、姫路市消防局の協力のもと、VR を用いた緊急走行シミュレータの研究開発を進めてきた[4-6]

(図 1)。本シミュレータを消防署に設置し、緊急走行経験の異なる消防士の運転データを収集することで、経験差に基づく運転行動の違いを分析してきた[4,5]。さらに、収集した運転データから生成した VR 緊急走行映像を用い、運転者自身による振り返り評価および他者による評価を実施し、消防士視点に基づくスキル評価の可能性を検討してきた[6]。

本研究では、これらの成果を一般運転者向けに拡張し、緊急車両の運転者視点を VR で体験できる「VR 緊急走行体験システム」を開発した。本システムでは、赤信号での交差点進入や渋滞時の反対車線走行など、緊急走行特有の状況を再現することで、一般運転者が緊急走行時の判断の難しさや心理的負荷を臨場的に理解し、避讓行動に対する意識向上と行動変容を促すことを目的とする。



図 1 消防士向け VR 緊急走行シミュレータ[4-6]



図 2 提案する VR 緊急体験システム

## 2. 提案システム

### 2.1 システムの概要

提案システムの外観を図 2 に示す。ヘッドマウントディスプレイ (HMD) として Meta Quest 3 を利用し、左右の Meta Quest 3 用コントローラ 2 つを 3 D プリンタで作成したアタッチメントで連結し、ステアリング型のコントローラとした(図 3)。アクセルおよびブレーキ操作は、左右のコントローラのボタンに割り当て、ステアリング操作はコントローラ全体の回転によって行う。

VR 都市環境の作成にあたり、3 次元都市モデルとして PLATEAU で提供される地理空間データを利用した。Unity 上に姫路市内の街並みを再現し(図 4)、周囲の車両・歩行者の飛び出しなどの挙動を設定することで、VR 空間内で緊急走行を体験できるようにしている。

### 2.2 VR 都市環境の構築

本節では、VR 都市環境の構築手順について述べる。まず、PLATEAU SDK for Unity を用いて、姫路市の 3 次元都市モデルを読み込んだ後、道路ネットワークを生成した。次に、PLATEAU SDK Toolkits for Unity で提供される Sandbox Toolkit を用いて、自車(緊急車両)以外の周辺車両・歩行者を配置するとともに、それらの移動経路であるトラックを作成することで、周辺車両・歩行者の行動を設定する。

Development of a VR Emergency Driving Experience System to Promote Drivers' Yielding Behavior toward Emergency Vehicles

†1 Yuki Nishiura, University of Hyogo

†2 Shingo Yamaguchi, Himeji Fire Bureau

†2 Akio Tanaka, Himeji Fire Bureau

†1 Hirotake Yamazoe, University of Hyogo



図 3 ステアリング型コントローラ



図 4 実装した VR 都市環境の街並みの例

Sandbox Toolkit は、信号制御を含む交通シミュレーションに対応しており、信号の変化タイミングを設定する。緊急走行を再現するため、周辺車両・歩行者については、緊急車両の接近に伴い移動するように設定している。また、PLATEAU で提供される都市モデルには、建物のみが含まれるため、臨場感を向上させるために Sandbox Toolkit を用いて街路樹を生成・配置した（図 4）。

また、後述の予備実験では利用していないが、Rendering Toolkit を用いることで、VR 都市環境の時間帯や天候を変化させることができるために、これを用いて、夜間や雨・雪といった低視認状況における緊急走行体験も実現できる。

### 2.3 緊急走行体験の実装

緊急走行体験の実装にあたっては、これまで実装してきた VR 緊急走行シミュレータ（図 1）を参考とした。緊急走行の臨場感を高めるため、サイレン音や赤色灯の光などを再現した。姫路西消防署において消防車のサイレンと無線交信を録音し、体験中に再生されるようにした。さらに交差点進入時にはモーターサイレンが自動的に流れよう実装した。

体験シナリオについては、共著者との相談の上、緊急車両の運転中に発生する、赤信号での交差点進入や渋滞時の反対車線の走行といった、緊急走行特有の状況を再現することとした。それを通じて、一般運転者が周囲車両の対応や心理的負荷を臨場的に体験できることを目指した。図 5 に体験シナリオを示す。姫路城周辺の赤矢印に沿って走行してもらうこととし、緑の○印で示す交差点では、車両・歩行者の飛び出しが発生し、前方の渋滞のため反対車線を走行してもらう。これらの体験により、一般運転者が身近な都市環境において、緊急走行の難しさを理解し、避讓行動の重要性を体感できることを目指す。

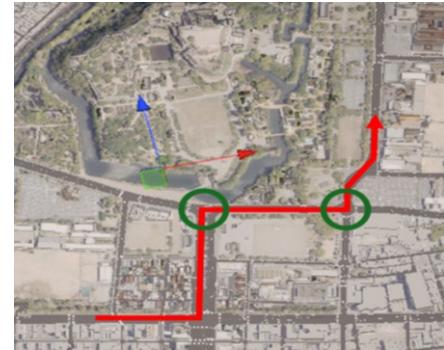


図 5 実装した緊急走行の体験シナリオ

## 3. 予備実験

提案システムは、2025 年 11 月 9 日に開催された姫路市防災フェアにおいて市民向けの展示を行った。その際、以下の実験条件・手順により、予備的な実験を行った。ここでは、その結果を報告する。

一般運転者が提案システムを用いて緊急走行を体験することで、避讓行動に対する理解や協力意識が向上するかを評価することを目的とする。ただし、展示における体験であるため、VR による緊急走行体験を 1 回のみ、アンケート評価も短時間で答えられるものとした。

### 3.1 予備実験の条件

予備実験では、体験中に緊急車両の走行が妨げられるイベントの有無による影響も評価するため、妨げあり条件、妨げなし条件の 2 条件を準備した。反対車線を走行する必要がある 2 か所の交差点（図 5 の緑の○印）において、妨げあり条件では対向車両が反対車線への進入を阻害し、一時的に前進できない状況を体験できるように設定した。

### 3.2 アンケート

緊急走行体験が一般運転者の避讓行動に与える影響を評価するため、体験後に表 1 に示すアンケートに回答してもらった。Q1～Q9 については、緊急走行体験と緊急走行に対する避讓行動に関する項目であり、1（まったくそう思わない）～7（とてもそう思う）の 7 段階リッカート尺度で回答してもらった。

Q1～Q5 については、計画的行動理論（Theory of Planned Behavior: TPB）に基づいて設定した項目である。計画的行動理論では、ある行動の実行は「行動意図」によって決定され、その行動意図は「態度」、「主観的規範」、「行動コントロール感」の 3 要因によって規定されるとする[7,8]。交通分野においては、運転行動分析に TPB が利用されており、例えば Yang らは都市交差点における譲り合い行動を、拡張 TPB モデルを用いて分析している[9]。本研究における Q1～Q5 は TPB に対応し、Q1（態度）、Q2（主観的規範）、Q3・Q4（行動コントロール感(PBC)）、Q5（行動意図）である。なお、PBC は「認知された力」と「コントロール信念」からなり、それぞれ Q3 と Q4 に対応する。

表 1 体験後アンケート項目

Q1	「道を譲る」行動は、望ましく、良い行動だと思いますか？
Q2	家族や他のドライバーなど多くの人は、「道を譲る」行動をすべきと考えていると思いますか？
Q3	「道を譲る」行動を、素早く安全に行う自信はありますか？
Q4	交通が混雑していても、「道を譲る」ことはどの程度難しい（又は簡単）と感じますか？
Q5	今後、「道を譲る」行動を、行うつもりですか？
Q6	今回の体験では、自分が運転する緊急車両の走行が妨げられたと感じましたか？
Q7	VR の中で実際に緊急車両に乗っているような感覚がありましたか？
Q8	今回の体験を通じて、緊急車両の運転者の立場や気持ちを理解できたと感じましたか？
Q9	この VR 体験を通じて、「道を譲ることの重要性」について理解が深まりましたか？
Q10	自動車の運転免許(普通免許)はお持ちですか？ [1.日常的に運転(経験: 年), 2.免許はあるがペーパードライバー, 3.免許無し]

Q6・Q7 は VR 走行体験に関する項目であり、Q6 は走行中に妨害を感じたかどうか、Q7 は VR 体験の臨場感を評価する。Q8 は緊急車両に対する共感、Q9 は体験を通じた学び・気づきに関する項目であり、また、Q10 では運転免許の有無、運転経験年数を尋ねた。

### 3.3 予備実験の流れ

まず、参加者に対してステアリング型コントローラの操作方法について説明した。その後、「妨げあり」または「妨げなし」のいずれか一方の条件で、VR による緊急走行体験を実施した。体験終了後には、上述のアンケートに回答してもらった。

## 4. 実験結果

予備実験には 7 名が参加し、そのうち 3 名は「妨害あり」条件、4 名は「妨害なし」条件を体験した。まず、走行中に妨害を感じたかどうか (Q6) のスコアを条件間で比較したところ、いずれの条件においても平均値が 6.0 であり、両条件間に大きな差は見られなかった。そのため、以降の分析では両条件の結果をまとめて分析する。

図 6 に各質問項目 (Q1～Q9) の平均スコアと標準偏差を示す。Q1～Q5 は TPB に基づく項目であり、避譲行動に対する態度 (Q1) は平均 6.9 と高く、参加者が「道を譲ることは望ましい行動である」と強く評価していることが確認された。主観的規範 (Q2) は平均 5.7 と比較的高く、他者から期待される行動としても認識されていた。

一方で、行動コントロール感 (PBC) を構成する、認知された力 (Q3) およびコントロール信念 (Q4) は、平均 4.1, 3.0

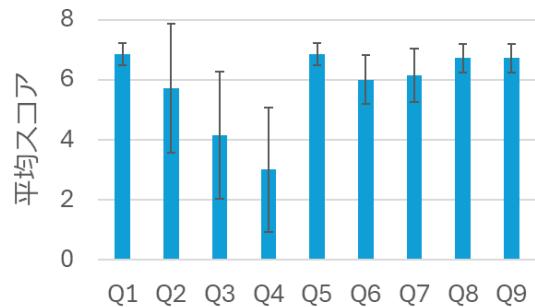


図 6 アンケート結果

と他項目より低かった。これは、参加者が避譲行動の必要性を理解している一方で、「実施できる自信」や「状況をコントロールできる感覚」が十分ではないことを示唆している。行動意図 (Q5) は平均 6.9 と高く、避譲行動をとる意欲自体は強いことが分かった。

VR 体験に関する項目では、臨場感 (Q7) は平均 6.1 であり、高い没入感を持って体験できていたことがわかる。また、共感 (Q8)、学び・気づき (Q9) はともに平均 6.7 といずれも高く、肯定的な反応が得られた。

全体として、TPB に基づく「態度」「主観的規範」「行動意図」は高く、VR 体験が共感や学び・気づきを促進する可能性が示された。一方で、PBC (Q3・Q4) が相対的に低かったことから、避譲行動の必要性は理解していても、その遂行には不安や難しさを感じていると考えられる。この結果は、避譲行動の必要性・重要性を伝えるだけでなく、具体的な行動手順の提示など、PBC を高める追加的な支援が必要であることを示唆している。

## 5. まとめ

本研究では、一般運転者が緊急車両の運転視点を VR で体験することで、緊急走行時における避譲行動への理解と協力意識の向上を目指した VR 緊急走行体験システムを開発した。提案システムは、PLATEAU の地理空間データを用いて姫路市内の都市環境を再現し、赤信号での交差点進入や渋滞時の反対車線走行など、緊急走行に特有の状況を体験できる。

予備実験の結果、体験者は避譲行動の必要性を理解し、共感や学び・気づきといった肯定的な反応を示した。一方で、行動コントロール感 (PBC) は相対的に低く、避譲行動の遂行に不安や難しさを感じている可能性が示唆された。ただし、本研究の結果は参加数が 7 名と少なく、「妨害あり／なし」条件の効果も明確には確認されておらず、本結果は予備的な知見に留まる。

今後は参加者数を増やした評価を行うとともに、妨害の感じ方に差が出るようなシナリオの設定を行い、システムの有効性を検証していく予定である。

## 謝辞

本研究の一部は、JSPS 科研費 23K26082 の助成を受けて実施した。

## 参考文献

- [1] 小池則満, 秀島栄三, 山本幸司, 地域特性と救急車の走行速度に関する分析－名古屋市を事例として－, 地域学研究, Vol. 30, No. 1, pp. 127–140, 1999
- [2] F. Herrera, J. Bailenson, E. Weisz, E. Ogle, J. Zaki, Building long-term empathy: A large-scale comparison of traditional and virtual reality perspective-taking, Plos One, Vol. 13, No. 10, 2018
- [3] M. Bujic, M. Salminen, J. Macey, J. Hamari, "Empathy machine": how virtual reality affects human rights attitudes, Internet Research, Vol. 30, No. 5, pp. 1407–1425, 2020
- [4] T. Sudou, S. Inoue, S. Yamaguchi, S. Nagata, H. Yamazoe, Preliminary Analysis of Emergency Vehicle Driving Behavior in Traffic Signal Violation Scenarios using a VR Simulator, The joint 34th International Conference on Artificial Reality and Telexistence & the 29th Eurographics Symposium on Virtual Environments (ICAT-EGVE 2024), 2024
- [5] 山添 大丈, 山口 晋吾, 永田 昌平, 緊急走行の事故防止教育に向けた VR 緊急走行シミュレータの開発と運転行動の分析, 安全・安心な生活と ICT 研究会 (ICTSSL) 2025 年度 5 月研究会, 2025
- [6] 小村太一, 井上草太, 山口晋吾, 永田昌平, 山添大丈, VR 緊急走行映像を用いた交通場面ごとの運転スキル評価と振り返りの予備検討, 第 30 回日本バーチャルリアリティ学会大会, 2025
- [7] I. Ajzen, The theory of planned behavior, Organizational behavior and human decision processes, Vol. 50, No. 2, pp. 179–211, 1991
- [8] I. Ajzen, Constructing a TPB questionnaire: Conceptual and methodological considerations, 2006.
- [9] B. Yang, K. Liang, X. Zhao, Xiaohua, L. Yang, Liping, W. Qu, Psychological influences on drivers' yielding behavior at the crosswalk of intersections, Cognition, Technology & Work, Vol. 22, No. 3, pp. 501–516, 2020

# Human-AI Teaming における AI の役割越境行動と信頼形成を検証するための実験系の構築

武安大樹<sup>†1</sup>鳥谷真佐子<sup>†1</sup>東京科学大学<sup>†1</sup>

## 1. はじめに

近年, AI 技術の発展に伴い, 人間と自律的 AI エージェントが共通の目標に向かって協働する「Human-AI Teaming (HAT)」が注目されている。HAT とは, 人間と AI がそれぞれ独立した役割を持ちながらも, 相互依存的に連携してタスクを遂行するチーム形態である[1]。従来の自動化研究では, AI は人間の指示に従って定められた手順を実行する“ツール”として扱われてきた。しかし近年の AI は, 動的な環境に適応し, 独立した分析・提案・意思決定を行うなど, 高い自律性を備えるようになり, 人間と共にチームの一員として機能する「チームメイト」へと変化しつつある。

このような背景から, HAT 研究は「AI がチームにどのように関わる, 人間との信頼関係をどのように形成・維持するのか」という新たな課題に注目している。先行研究では, HAT の導入がタスク効率を高める一方で, チーム内のコミュニケーションや協調, さらには信頼の低下を招く可能性が指摘されている[2]。また, 体系的レビューによれば, HAT における信頼研究はエージェントの信頼性 (reliability) や透明性 (transparency) などのパフォーマンス要因に偏り, AI によるコミュニケーションスタイルや信頼修復行動といった行動的要因 (behavior-based factors) の実証的検討が不足していることが報告されている[3]。こうした研究の偏りは, HAT における信頼の理解を限定的なものにしており, AI が社会的・行動的にどのように信頼を形成・修復するかという観点の理論的整理を妨げている。

したがって, 今後の HAT 研究では, AI を単なるツールではなく「協働主体」として捉え, 人間と AI の相互依存関係を前提に, 協働や信頼の新たなメカニズムを探求することが求められる。

## 2. 目的

本研究ではこの潮流の中で, AI が自律的に行動し, チームに介入するような状況に着目する。特に, AI が自身に割り当てられた役割を超えて, 人間の作業を「良かれと思って」肩代わりするような過剰支援 (over-support) の状況に焦点を当てる。このような行動は, 一見するとチーム全体の効率を高めるように見えるが, 同時に人間の主体性や役割認識,

さらには AI への信頼に負の影響を及ぼす可能性がある。実際, AI による能動的支援が利用者の能力に対する自尊心を損ない、システムへの満足度を低下させる事例も報告されている[4]。

本研究はこのような観点から, AI の過剰支援がチームの協働過程や信頼形成にどのような影響を与えるのかを明らかにし, 人間と AI が文化的・社会的文脈の中でどのように共進化しうるかを探ることを目的とする。

## 3. 方法

チームでの業務はプロジェクト型とルーチン型に分けられる。本研究では, より不確実性が高くバックアップ行動が頻繁にみられるプロジェクト型業務を想定した。

チームプロセス研究において, メンバー間のタスク遂行を支援する行動としてバックアップ行動が広く知られている。バックアップ行動には, 口頭でのフィードバックやコーチングといった言語的支援に加え, 行動レベルでの直接的なサポート, さらには他者のタスクを代わりに引き受けける行動が含まれる[5]。本研究では, AI エージェントおよび参加者に発話を行わせない設計とした。発話を用いると, 指示・助言・フィードバックといった言語的支援が生じ, チーム内に指導者と被指導者という階層的関係が生まれる可能性がある。こうした上下関係や指摘の関係性は実験に過剰な複雑性をもたらすため, 本研究ではこれを排除し, 対等な関係性における行動レベルのバックアップに焦点を絞った。

本研究では, 口頭でのフィードバックやコーチングといった言語的支援を除くバックアップ行動に加えて, AI 特有のバックアップ行動を想定した。すなわち, AI は一貫した基準での評価能力や修正能力により, 人間が完了したタスクを無機質に修正しうるため, 完了済みタスクの修正行為も新たなバックアップ行動に含まれると考えた。以上をふまえ, 本研究ではこれらのバックアップ行動を実験的に再現するタスク環境の構築を試みた。

本研究では, 4 人 1 チームの塗り絵タスクを用い, そのうち 1 名を AI エージェント (Wizard of Oz 方式) として設定する。塗り絵を採用したのは, AI の介入が視覚的に把握しやすく, 絵を四象限に分ける構造により参加者の担当領域が明確になる一方で, 1 枚の絵が象限をまたいで成立するため, 領域を越えた自然な介入が生じる環境を構築できるためである。また, 塗り絵タスクは色の配置, 塗り残し, 他メンバ

Role-Boundary-Crossing Behaviors of AI and Trust Formation in Human-AI Teaming

<sup>†1</sup> Daiki Takeyasu, Institute of Science Tokyo

<sup>†1</sup> Masako Toriya, Institute of Science Tokyo

一の進捗状況などの視覚的判断を伴う複雑なタスクであり、実験条件に応じた適切な介入を自律的に実行できるAIシステムの実装は現時点では困難である。そのため、WoZ方式を採用することで、実験デザインに必要な介入行動を確実に実行しつつ、条件間での統制を厳密に保つことを可能にした。図1に塗り絵タスクの例を示す。



図1. 塗り絵タスク

本研究で設計するAIによる介入は、補完的介入、先塗り介入、上塗り介入の三種類である。いずれも現実のチームでみられる役割越境行動を模すことを想定している。補完的介入は、チームメンバーが担当領域を途中まで行っている作業に倣ってその領域の塗り絵作業を行う行動であり、「行動レベルでの直接的なサポート」に対応する。これは、現実のチームにおいては他者の未達業務を補う行動にあたる。先塗り介入は、チームメンバーの未着手の担当領域を先に塗る行動であり、「タスクを代わりに引き受ける行動」に相当する。これは、チームにおいて他者の担当タスクを先行して着手する支援にあたる。上塗り介入は、チームメンバーの塗り絵作業が完了している領域を修正する行動であり、「完了済みタスクの修正行為」にあたる。これは、成果物全体の整合性を保つために他者により行われた業務を修正するサポートである。

表1にこれらの行動とバックアップ行動の対応を示す。

表1. バックアップ行動と塗り絵タスク介入の対応

バックアップ行動	塗り絵タスクでの介入
行動レベルでの直接的なサポート	補完的介入
タスクを代わりに引き受ける行動	先塗り介入
修正するサポート	上塗り介入

これらのAIによる行動は、状況によってはチームの調整プロセスやメンバーの受け止め方に影響を及ぼす可能性がある。本研究のデザインは、協働における役割越境の特性を再現しつつ、AIによる介入がチームダイナミクスにどのように

作用するかを実験的に検証する点に特徴がある。

本実験に先立ち、まず、提案する実験タスクが実際のプロジェクト型業務における協働の特性を適切に反映しているかを検証するため、Webアンケートおよび動画刺激を用いた予備調査を行う。予備調査では、実務経験を持つ社会人を対象に、実際の業務におけるこれら3種類のバックアップ行動（補完的介入・先塗り介入・上塗り介入に相当する行動）の発生頻度や、それに対する感情的反応を聴取する。あわせて、塗り絵タスクにおける各介入動作の動画を提示し、それらが意図したバックアップ行動として認知されるか、また実際の業務との類似性を感じるかを確認する。この調査によってタスクの生態学的妥当性が確認された後、実際に参加者を募った対面実験を行う。そこでは、AIによる過剰支援（役割越境行動）に対する人間の感情や信頼の変容プロセスが、対人間の場合と比較してどのように異なるのかを検証する予定である。

#### 4. 期待される成果・意義

本研究の学術的貢献は、主に以下の2点にある。第1に、プロジェクト型業務のメタファーとしての「塗り絵タスク」の提案である。これにより、従来の単純作業タスクでは観察が困難であった「役割越境」や「過剰支援」に対する微細な心理変容を、実験環境において再現・観察可能にする。第2に、AIの自律性設計への示唆である。AIが良かれと思って行う介入（過剰支援）が、かえって人間の信頼や主体性を損なうプロセスを解明することで、今後のHAT（Human-AI Teaming）設計における「支援の適正範囲」や「文化的・社会的文脈に応じた振る舞い」に関する重要な理論的基盤を提供できる。

#### 5. 今後の課題

直近の課題は、Webアンケートおよび動画刺激を用いた予備調査を通じて、提案タスクの妥当性を確立することである。その後、Wizard of Oz法を用いた本実験を実施し、3種類の介入（補完・先塗り・上塗り）が参加者の感情や対AI信頼感に与える影響を定量・定性の両面から分析する。特に、介入者が「人間」である場合と「AI」である場合の受容性の差異に着目し、AI特有の信頼毀損・修復メカニズムのモデル化を目指す。

#### 参考文献

- [1] O'Neill, T. A. et al., "Human-Autonomy Teaming: A Review and Analysis of the Empirical Literature," *Human Factors*, 2020. <https://doi.org/10.1177/0018720819894214>.
- [2] Schmutz, J. B. et al., "AI-teaming: Redefining Collaboration in the Digital Era," *Curr. Opin. Psychol.*, 2024. <https://doi.org/10.1016/j.copsyc.2024.101837>.
- [3] Duan, W. et al., "Trusting Autonomous Teammates in Human-AI Teams," *Proc. CHI '25*, ACM, 2025. <https://doi.org/10.1145/3706598.3713527>.
- [4] Kim, S. & Song, H., "When AI-Based Agents Are Proactive: Implications for Competence and System Satisfaction in Human-AI Collaboration," *Int. J. Hum.-Comput. Interact.*, 2025. <https://doi.org/10.1080/10447318.2023.2289947>.
- [5] Marks, A. M., Mathieu, J. E., & Zaccaro, S. J., "A Temporally Based Framework and Taxonomy of Team Processes," *Acad. Manag. Rev.*, 2001. <https://doi.org/10.5465/amr.2001.4845785>

# 可搬型 IoT を活用した多人数対象昇降・転倒初動検知に向けた IMU データ準備時間短縮システム

小島涼太郎<sup>†1</sup>東京電機大学<sup>†1</sup>岩井将行<sup>†2</sup>東京電機大学<sup>†2</sup>

## 1. はじめに

高所作業を伴う現場では、脚立や立ち馬を用いた昇降動作が日常的に行われており、建設業・設備点検・道路維持管理などの領域では小規模な昇降を繰り返す作業が多い。その一方で、令和 3 年の労働災害による死者者数のうち建設業が 288 人を占めており、事故の型別では高所からの「墜落・転落」が 110 人と最も多いことから、昇降時の姿勢変化や重心移動の乱れが重大事故に直結しやすいことが示されている [1]。特に、昇降動作中は片手で工具や資材を保持する場面も多く、不安定な姿勢になりやすいことから、危険兆候を事前に把握する仕組みが求められている。

近年、小型 IoT デバイスを用いて加速度データを取得し、転倒検知や異常動作分類を行う研究も行われている [2] [3][4][5]。しかし、従来研究の多くは「転倒後の検知」や「急激な姿勢変化の判定」といった事後的な異常検知に焦点があり、昇降動作中の危険兆候を早期に検知するための基盤整備や、IMU データ取得プロセスの効率化については十分に検討されてこなかった。

実際の現場で動作計測システムを普及させるためには、

- ・デバイスセットアップの容易さ
- ・データ取得～機械学習準備までの運用時間の短縮
- ・日常的に使いやすいシンプルなワークフロー

が重要となる。複雑なデータ取得手順は作業者の負担となり、運用現場への定着を妨げる可能性がある。

また、文献[6] の先行研究では、データ取得からモデルの生成までの流れ自体は構築されていたものの、SD カードを用いたデータ取得方式では、小型 IMU デバイス内の SD カードに一度データを保存し、後から PC に取り出して処理する必要があるなど、データ収集からモデル生成までの準備に時間を要する点が現場での実装を妨げる要因となっていた。

本研究では、昇降動作中に生じる危険兆候を早期に検知するために小型 IMU デバイスを用いて昇降動作データを取得し、Node-RED を介した即時データファイル生成プロセスによりデータ取得基盤を構築する。これにより、小型 IMU

デバイスから取得したデータをリアルタイムに処理し、モデルの構築に必要なデータ準備を効率化する仕組みを実現している。さらに、先行研究と比較することで、データ取得からモデル生成への準備までに要する運用時間をどの程度削減できるかを検証し、現場運用に適したデータ収集フローを提案した。

## 2. 関連研究

### 先行研究との差異

先行研究では、小型 IMU デバイスをヘルメットに装着し、小型 IMU デバイスに搭載した SD カードに加速度データを一度保存する。その後 SD カード内のデータを 1 行ずつ読み出して MQTT 通信を用いて Node-RED を介し Raspberry Pi に転送し、ファイルとして保存する方式を採用していた。また、生成されたデータファイルを Edge Impulse にアップロードする際には、Raspberry Pi から scp コマンドを用いて PC へファイルを転送する必要があり、運用上の手間が大きいという課題が存在していた。さらに、先行研究で使用されていた M5Core2 は機能性は高いものの筐体が比較的大きく、作業者の身体に取り付けて運用する場合には装着負担が大きいという問題も残されていた。

これに対し本研究では、小型 IMU デバイスを胸に装着し、小型 IMU デバイスとしてより小型・軽量な M5ATOMS3 を採用することで、身体への装着負担を低減し、現場での実運用に適した形で計測デバイスの携行性・利便性を向上させた。またデータ処理方式においても、取得したデータを SD カードを使用せずに MQTT 通信を用いて Node-RED を介して Raspberry Pi へ直接送信し、即時にファイルを生成する方式を導入した。さらに Raspberry Pi に Samba を導入することで、PC から Raspberry Pi 内のファイルへ直接アクセスできる環境を整備し、Edge Impulse へのアップロードに際してファイル転送作業を不要とした。

これにより、小型 IMU デバイスの携行性向上による 現場適応性の改善に加えて、データ保存に要する時間の短縮と、モデル学習データ準備における 作業効率の大幅な向上 が期待できる。

How to Prepare Your BTI Symposium Reports in MS-Word

†1 RYOTARO KOJIMA, Tokyo Denki University

†2 MASAYUKI IWAI, Tokyo Denki University

### 3. 提案システム

#### 3.1 提案システムの概要

提案システムの概要を図 1 に、Node-RED フローを図 2 に、使用するデバイスを図 3 に示す。

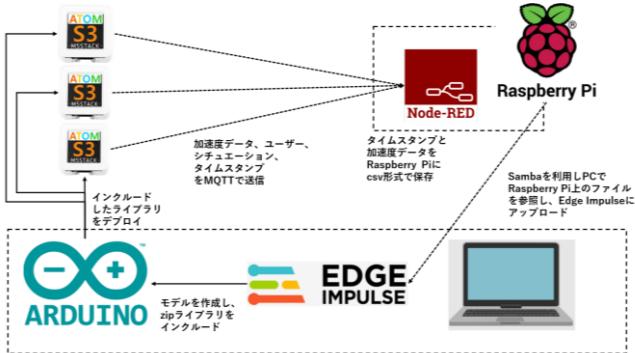


図 1 システム構成図

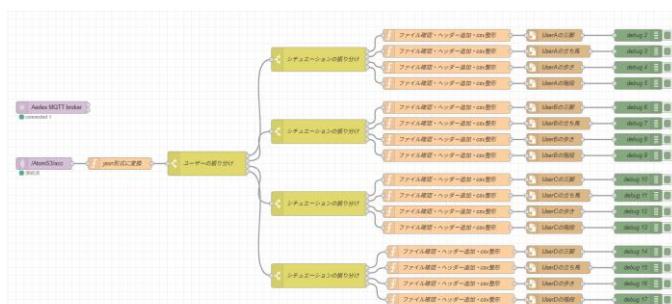


図 2 Node-RED フロー図

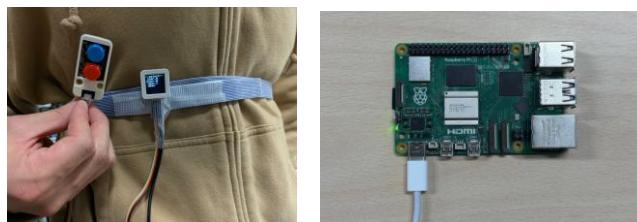


図 3 使用するデバイス

(左 : M5ATOMS3 ・ デュアルボタン  
右 : Node-RED の Raspberry Pi5)

動作計測は先行研究と同様に、被験者は M5ATOMS3 に接続されたデュアルボタンを片手に持ち、大学の廊下を一往復歩き、階段・三脚・立ち馬の上り動作をそれぞれ 10 回ずつ行った。M5ATOMS3 は 16 ミリ秒間隔で加速度データを取得し、取得した加速度データは、ユーザー名、シチュエーション、16 ミリ秒ごとのタイムスタンプを 1 つのデータパケットとしてリアルタイムに CSV 形式のデータとして MQTT 通信を介し、Node-RED を経由して Raspberry Pi に保存する。Node-RED では、受信した CSV 形式のデータを switch ノードで振り分けやすくするために、一度 JSON 形式へ変換する。switch ノードでユーザーおよびシチュエーションごとに分岐させ、タイムスタンプおよび加速度データのみを抽出・CSV に整形し、「ユーザー名\_シチュエーション.csv」という形式のファイル名で Raspberry Pi に保存する。Raspberry Pi には Samba を導入しており、PC から Raspberry Pi に保存された CSV ファイルに即時アクセスが可能であるため、ファイル転送処理を行うことなく Edge Impulse へアップロード可能である。

Edge Impulse では、特徴量抽出に Raw Data と Spectral Analysis、学習ブロックに Classification を用いることで、ユーザー個人に適応した動作識別モデルを生成する。生成したモデルは Arduino Library としてデプロイし、Arduino IDE を用いて M5ATOMS3 に書き込むことで、小型 IMU デバイス自体に個人専用の動作識別モデルを搭載することができる。

#### 3.2 システムの特徴と利点

提案システムの主な特徴は以下の通りである。

- ・ 現場での簡易操作：電源投入後すぐにデータ収集が可能であり、複雑なセンサ準備や SD カード管理を必要としない。
- ・ 複数作業者の同時データ取得：ユーザー選択やモード切り替えをデバイス上で直感的に操作でき、複数作業者からデータを同時に取得することで、データの収集時間を短縮することができる。
- ・ データ収集からモデル生成までの効率化：ネットワーク経由で自動取得されたデータは、PC に保存することなく、Samba を利用してそのまま Edge Impulse にアップロードでき、個人適応モデルの構築を容易にする。

これにより、提案システムは現場での負担を最小限に抑えつつ、昇降動作に潜む危険兆候を早期に把握するための個人適応型データ収集基盤として有効である。

### 4. 提案システムの評価

先行研究の手法を用いて測定された動作計測からモデル生成までの時間と本研究で提案するシステムの動作計測からモデル生成までの時間を比較し、評価する。

実験の様子を図 4 に示す。



図4 実験の様子

先行研究と提案システムの処理時間の比較を表1および表2に示す。

	被験者 A	被験者 B	被験者 C
動作計測	11分53秒	12分37秒	11分01秒
データファイルセット	9分43秒	8分57秒	8分26秒
モデル生成	10分01秒	11分25秒	9分07秒
合計	31分37秒	32分59秒	28分34秒

表1 先行研究に基づく処理時間の計測結果

	被験者 A	被験者 B	被験者 C
動作計測	10分54秒	12分37秒	11分53秒
データファイルセット	0秒	0秒	0秒
モデル生成	10分41秒	11分18秒	9分45秒
合計	21分35秒	23分55秒	21分38秒

表2 本研究の処理時間の計測結果

先行研究では動作計測後に SD カードに書き込まれたデータを読み出し、Node-RED を介して Raspberry Pi へ保存する必要があり、この工程に 8~9 分を要していた。一方、本研究ではデータをリアルタイムに Node-RED を介して Raspberry Pi へ保存し、手動による転送作業を必要としないため、本研究に要する時間は 0 秒となる。

表1および表2より、本研究の提案システムは先行研究と比較して、動作計測からモデル生成までに要する総処理時間を大幅に短縮できていることが確認できる。特に、先行研究では SD カードの読み出しおよびファイル保存に要していた時間を、本研究では完全に排除することで、処理時間を 0 秒とすることに成功した。

被験者ごとの総処理時間の比較を行うと、

- ・被験者 A : 31分37秒 → 21分35秒 (31.7%短縮)
- ・被験者 B : 32分59秒 → 23分55秒 (27.5%短縮)
- ・被験者 C : 28分34秒 → 21分38秒 (24.3%短縮)

となり、最小で 24.3%、最大で 31.7% の処理時間短縮が達成されている。

さらに、3名の平均値で比較すると、

- ・先行研究 : 31分10秒
- ・本研究 : 22分23秒

となり、平均した時間で比較すると 28% の削減が実現された。

これは、データ転送作業を不要としたリアルタイム保存方式が、システム全体のワークフローを大きく効率化していることを示す結果である。

また以下に Edge Impulse にアップロードしたデータに基づいて作成したモデルの精度を図5に示す。

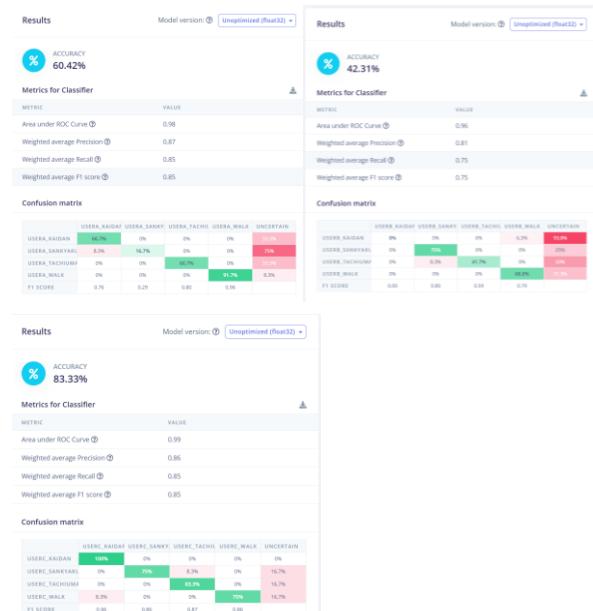


図5 生成したモデルの精度

図5に示すように、生成されたモデルの分類精度は被験者によって大きく異なり、安定して高い精度を得るには至らなかつた。また、精度が高く出なかつた要因として、被験者ごとの歩行速度や揺れ方の違いによって、ウィンドウ内の加速度の動き方が安定せず、特徴量がばらついた可能性

がある。特に、歩行のテンポや身体の動きが一定でない場合には、1 ウィンドウ内に複数の動きが混在しやすく、特徴抽出がさらに不安定になったと考えられる。

また、クラス間のデータ量の偏りや、設定したウィンドウサイズ (1500 ミリ秒) とストライド (500 ミリ秒) が動作周期と適合していなかった可能性もあり、これらがモデル学習に影響を与えた可能性も考えられる。

## 5. まとめ

本研究では、小型 IMU デバイスから取得した加速度データを、SD カードを介さずにリアルタイムで Raspberry Pi に保存し、Edge Impulse へアップロードしてモデル生成まで行う一連の処理フローを構築した。本研究で提案したシステムは、先行研究で課題となっていた ファイル転送作業の手間と時間的コストを解消し、動作計測からモデル生成までのプロセス全体を効率化することを目的としている。

先行研究と本研究の処理時間を比較した結果、最も大きな違いは SD カードから Raspberry Pi へのファイル転送工程の有無であった。先行研究ではこの工程に 8~9 分 程度を要していたのに対し、本研究のシステムではデータを計測と同時に Raspberry Pi へ保存するため、転送時間は 0 秒 となつた。

また、動作計測からモデル生成までの合計時間を比較すると、先行研究では 約 28~32 分 を要していたのに対し、本研究では 約 21~23 分 に短縮された。

これは平均して 28% の処理時間削減 に相当し、動作計測からモデル生成までのワークフロー効率が大幅に向上したことを見ている。

さらに、本研究のシステムはリアルタイム保存と Samba を通じて Raspberry Pi 内のデータを PC から即時に参照できるため、データ管理の手間を大幅に削減できる。これにより、操作ミスの防止やデータの一貫性確保にもつながり、より効率的で再現性の高い実験運用が可能となる。

また本研究の提案システムを用いて生成した動作分類モデルについては、全体として十分に高い精度を得るには至らず、被験者間および動作クラス間で一定のばらつきが確認された。これは、ウィンドウサイズとストライドの設定、被験者ごとの加速度特性の違い、データ量の偏りといった要因が考えられる。しかし、これらの要因は適切なパラメータ調整によって改善可能であり、本研究のモデル性能にはまだ向上の余地が残されている。今後、ウィンドウ長の最適化、収集するデータの量、個人適応型モデルの検討などを行うことで、より安定した分類性能が期待される。

以上より、提案したシステムは先行研究に比べて効率性・操作性・信頼性を向上させるものであり、行動認識モデルの迅速な構築が求められる実応用において有用であると考えられる。

## 6. 参考文献

- [1] 厚生労働省 第 14 次労働災害防止計画  
<https://www.mhlw.go.jp/content/11200000/001116307.pdf>
- [2] 五艘隆志,武藤一伸,機械学習を用いた建設現場の労働環境・生産性データ収集分析システム構築に用いる各種端末開発, 助成研究報告書, 2021-2025.
- [3] 魚谷 義也,葉 臣,山本 幸平,大槻 知明:加速度センサを用いた LSTM に基づく転倒検知, 電子情報通信学会, 2021 年 2 月
- [4] 和田 康寛,岩井将行:無線 9 軸センサノードを用いた機械学習アルゴリズムによる睡眠状態検知システム, 電子情報通信学会, 2021 年 10 月
- [5] 大井健矢, 松下光範, 堀寛史, 高岡良行. ウェアラブルセンサによる歩行データからの転倒動作検出の試み. 人工知能学会 インタラクティブ 情報アクセスと可視化マイニング研究会 (第 28 回), 2022 年 3 月
- [6] 高橋 大和,岩井 将行,高所作業員向け IoT ヘルメットのためのユーザ個別の 6 軸 IMU センサの値に基づくモデル生成時間短縮システム, 情報処理学会第 86 回全国大会, 2024 年 3 月発表予定

# 上半身姿勢情報を用いた作業専念度の リアルタイム推定モデルとアバタの表情表示連携

山口美佳 <sup>†1</sup>  
東京電機大学      岩井将行 <sup>†2</sup>  
東京電機大学

## 1. はじめに

学習や作業支援の分野では、ユーザの集中状態を客観的に把握し、適切にフィードバックすることが重要である。従来の研究では、脳波、心拍、視線といった生体信号が集中度の推定に利用されてきた。しかし、これらは装置の装着による負担や導入コスト、個人識別性によるプライバシ問題が指摘されている。

一方、身体の姿勢は認知的な負荷や作業への没頭度と関連することが報告されている。非接触で計測できる姿勢センシングは、日常的な環境でも導入しやすい手法として注目されている。近年は OpenPose や MediaPipe 等の技術により、単眼カメラから骨格の推定が可能となり、映像を用いた学習者の分析が進展している。

しかし、推定した作業専念度を数値やグラフとして直接提示すると「評価・監視されている」という感覚を与え、逆効果となることが教育心理学の観点から指摘されている。そのため本研究では、個人を特定せず集団全体の傾向として可視化する手法を採用し、心理的な負担を抑えたインターフェース設計を目指す。

## 2. 関連研究

### 2.1. 身体動作・姿勢を用いた推定手法

生体信号を用いない手法として、身体の動作や姿勢から集中度を推定する研究もおこなわれている [1, 2, 3]。例えば、鶴岡ら [4] は圧力センサを搭載したクッション型デバイスを用い、座面にかかる圧力分布の変化から姿勢を検出し、デスクワーク時の集中度推定を行っている。

このようなクッション型デバイスを用いた手法は、身体に機器を直接装着する必要がない「非拘束型」であるため、ウェアラブルセンサに比べてユーザの身体的負担が少ないという利点がある。しかし、利用には専用のハードウェア（センサ内蔵クッション）を作業環境に設置する必要があり、導入コストや環境への依存性が課題となる。

これに対し本研究では、PC に付属する Web カメラ等の単眼カメラのみを用いる手法を提案する。この手法は、ユーザへの接触が一切ない「非接触型」であり、かつ特殊なデバイスを追加で用意する必要がないため、より日常的な学習・作業環境へ容易に導入できる点が特徴である。

### 2.2. 本研究の位置づけ

前節で述べたセンシングの課題に加え、学習・作業支援においてはフィードバックの方法も重要である。個人の推定結果を数値として直接提示することは、ユーザに「監視・評価されている」という心理的負担を与え、かえってパフォーマンスを低下させるリスクがある。

そこで本研究では、以下の 2 点のアプローチにより、日常的な環境で導入しやすく、かつ心理的負担の少ない支援システムの実現を目指す。

1. **単眼カメラによる非接触センシング**：特殊なデバイスを必要とせず、PC 内蔵カメラ等から得られる映像のみを用いて、MediaPipe Pose により上半身の姿勢情報を取得する。これにより、物理的な制約を最小限に抑え導入が可能となる。
2. **アバタを用いた作業専念度の可視化**：推定結果を数値として直接表示するのではなく、アバタの表情（リラックス・通常・集中）という、全体の雰囲気として抽象化してフィードバックを行う。特に集団で利用する場合には、個人の値を直接表示せず全体の平均値を算出してアバタで提示する。これにより、情報の解像度を意図的に下げ、個人を特定しない形でプライバシを保護しつつ、監視感や束縛感を与えずに学習環境の状態を共有する。

## 3. 提案手法

本研究では、単眼カメラから取得した上半身の姿勢情報を用いて、作業者の作業専念度（以下、専念度）をリアルタイムに推定し、複数人の平均専念度をアバタ表情として提示する仕組みを構築した。提案手法は、(1) 骨格推定、(2) 特徴量抽出、(3) 専念度推定モデル、(4) フィードバック手

<sup>†1</sup> MIKA YAMAGUCHI, Tokyo Denki University  
<sup>†2</sup> MASAYUKI IWAI, Tokyo Denki University

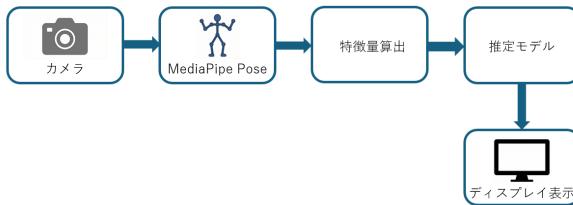


図1 提案システムの処理フロー

法の4段階で構成される。

### 3.1. 骨格情報の取得

MediaPipe Pose を用いて、頭部・肩部・鼻を中心とした上半身の骨格 21 点を 1080p/30fps の映像から推定した。取得された座標系列に対し、15 秒を 1 ウィンドウとする時間窓を設定し、各ウィンドウに含まれる姿勢変化を基に、姿勢の安定性・動きの大きさ・傾きなどを計算する。

### 3.2. 特徴量抽出

取得された骨格座標系列に対し、15 秒間のウィンドウを設定し、各ウィンドウから以下の 5 つの特徴量を算出した。これらは、作業への集中に伴う「静止状態」や、散漫時の「動き・姿勢の崩れ」を捉るために選定した。

- **身体動作量（平均・標準偏差）**：前後フレーム間における各骨格点の位置差分の平均値および標準偏差。平均値は身体全体の動きの大きさを、標準偏差は動作のばらつきを表す指標として用いる。
- **頭部の沈み込み**：両肩の中点から鼻までの垂直距離の平均値。集中時の前傾姿勢や、覚醒レベルの低下（居眠り等）による頭部の垂れ下がりを検出する。
- **肩の左右非対称性**：両肩の高さの差分平均。頬杖をつく、あるいは体が左右に傾くといった姿勢の崩れを反映する。
- **上体の傾き**：両肩を結ぶ直線の角度の平均値。垂直軸に対する上半身の全体的な傾斜傾向を捉える。

### 3.3. 専念度推定モデル

算出された特徴量を 1 つのベクトルとしてまとめ、ランダムフォレスト回帰モデルによる学習を行った。学習データの構築にあたり、被験者 3 名を対象にデータ収集実験を実施した。多様な集中状態を取得するため、「制限時間付きタイピング（高集中）」および「動画視聴（リラックス）」の 2 種類のタスクを設定した。ラベル（専念度 0 ~ 10）の付与には、実験終了後に被験者自身が録画映像を見返しながら主観評価を行う回顧法を採用した。これにより構築されたモデルは、映像入力に対して一定時間ごとに専念度を出

力でき、リアルタイムな処理が可能である。

### 3.4. フィードバック手法

専念度の提示においては、個人値を直接表示すると、「監視・評価されている」といった心理的負担を与える可能性がある。単独の利用であれば、自己管理の指標として利用することに一定の有用性がある。しかし、複数人が存在する環境では、個人の数値が他者に露呈することへの抵抗感が強く、かえって学習・作業意欲をそぐ恐れがある。

そこで本研究では、以下の方針に基づいてフィードバックを設計した。

- 集団利用における個人の専念度は**非表示**とする
- 複数人が検出された場合、専念度の平均値のみをアバタの表情として提示する

専念度と表情の対応関係は以下の通りである。

表1 専念度と表情の対応

専念度	表情
0~3	リラックス
4~7	通常
8~10	集中



図2 アバタの表情変化（左から：リラックス、通常、集中）

数値データが「個人の能力評価」として受け取られやすいのに対し、アバタ表示は「場の空気」として直感的に認識されやすい。そのため、相互監視からのプレッシャーから解放され、気負わずに自分の作業に向き合うことが可能となる。

## 4. 実装

本研究では、構築した専念度推定モデルをリアルタイムに動作させるシステムを実装した。このシステムは、(1) カメラ映像の取得、(2) 骨格推定、(3) 専念度推定、(4) アバタ表情表示の 4 つの処理から構成される。実装には Python を使用し、骨格推定ライブラリとして MediaPipe Pose を用いた。

#### 4.1. 处理フロー

映像を MediaPipe Pose に入力し, 上半身の骨格情報をフレームごとに取得する. 取得された骨格データはバッファに蓄積され, 直近 15 秒のデータを対象に特徴量を計算する. その後, ランダムフォレスト回帰モデルにより専念度が算出される.

専念度の微細な変動によるちらつきを防ぐため, 推定処理およびアバタの表情更新は 1 秒間隔で実行する. これにより, 安定した連続的な専念度の変化を取得可能にしている.

#### 4.2. アバタ表情の表示と複数人対応

本システムは MediaPipe Pose を用いて画面内に存在する人物を自動的に検出し, それぞれの骨格から個別の専念度を算出する. フィードバックにおいては, 前節で述べたプライバシ保護の方針に従い, 個人の推定値は表示しない. 代わりに, 検出された全ユーザの専念度の平均値を算出し, その値に応じた表情 (リラックス・通常・集中) をアバタに反映させる.

この仕組みにより, 個人を特定する監視感を与えることなく, 集団全体の専念度を直感的に把握可能なインターフェースを実現した.

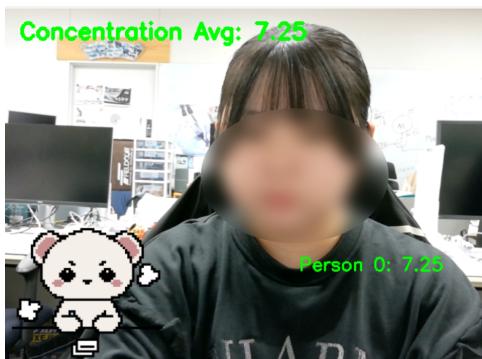


図 3 実装システムの動作画面 (単独)

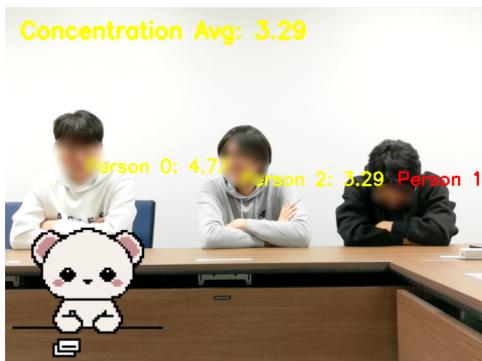


図 4 複数人同時検出時の動作画面

#### 5. 評価

構築した専念度推定モデルの精度検証を行うため, 収集したデータを用いた評価実験を行った.

##### 5.1. 実験設定

3 名の被験者によるデータセット (全データからランダムに 20% を抽出) に対し, モデルによる推定値と, 被験者がつけた正解ラベルとの誤差 (差の絶対値) を確認した.

##### 5.2. 実験結果

結果として, 正解ラベルが中間領域 (3.5 ~ 7.5) のデータに対しては, 誤差が平均して 1.5 未満となり, 高い精度で推定が可能であった. 一方で, 正解ラベルが極端に低い (0 ~ 3) 場合や高い (8 ~ 10) 場合においては, モデルの出力が中央値 (5 付近) に寄る傾向が見られ, 誤差が大きくなるケースが確認された.

##### 5.3. 考察

実験結果より, 本モデルには出力値が平均付近に集まる中心化傾向が見られた. これは, 学習データにおいて極端な集中状態や完全なリラックス状態のサンプル数が相対的に不足していたことが原因と考えられる.

しかし, 特徴的な姿勢が見られる際には適切に「集中」や「リラックス」の判定が行われており, システムとして一定の実用性は保たれているといえる.

今後は, 極端な集中・散漫状態のデータを重点的に収集し, データの不均衡を解消することで, 全体の推定精度を向上させる必要がある.

#### 6. おわりに

本研究では, 単眼カメラから得られる上半身姿勢情報を用いて専念度をリアルタイムに推定するモデルを構築し, 複数人の平均専念度をアバタ表情として可視化する仕組みを提案した. 非接触で計測可能な姿勢情報を利用することで, 日常的な環境でも導入しやすく, 心理的負担の少ないフィードバックが可能となる.

本システムでは, 骨格推定, 専念度推定, アバタ表情表示を連続的に実行し, 集団状態を直感的に把握できる可視化を実現した. とくに, 個人値を提示せず平均値のみを表情に反映する点は, 専念度の直接提示が与える監視感を抑えつつ, 学習環境や作業環境の状態を柔らかく共有できる点で有用である.

今後はデータ量の拡充や特徴量の改善による推定精度向上に加えて, 工事現場の安全研修のような高い集中維持が不可欠な場面や, 実演などで学習者を直視できない場面な

ど、アバタによる直感的な把握が指導の補助となる環境での有用性を検証したい。また、リアルタイム性や複数人同時推定の安定性向上を図り、実運用に向けたシステムとして発展させていく予定である。

## 参考文献

- [1] 亀井諭. 学習時の姿勢と行動の計測による集中度合いの推定. Master's thesis, 中央大学大学院 理工学研究科 経営システム工学専攻.
- [2] 立花優斗, 今井順一. e-learning 学習者の上半身姿勢情報を利用した集中度推定. 第 15 回情報科学技術フォーラム (FIT2016), No. N-018, 2016.
- [3] 清野悠希, 佐藤哲司. 重心動搖と姿勢計測による学習状態の推定. 第 7 回データ工学と情報マネジメントに関するフォーラム (DEIM2015), No. D7-1, 2015.
- [4] 鶴岡秀樹, 後藤隆太郎, 横田悠右, 成瀬康, 矢入郁子. クッション型デバイスによる着座姿勢と学習集中度の評価. 第 31 回人工知能学会全国大会論文集, No. 3N2-3, 2017.

## 複数の AIoT アバタを活用した情報提供システムの構築

木田勇輝<sup>†1</sup> 武藤英樹<sup>†1</sup> 千葉俊彦<sup>†1</sup> 小木哲朗<sup>†1</sup>  
慶應義塾大学<sup>†1</sup>

### 1. はじめに

筆者らはこれまでに、実空間に存在する様々なモノに小型 IoT デバイスを取り付け、遠隔ユーザがそれを操作しそのモノとして現場の人とコミュニケーションを行うことが可能な IoT アバタを開発してきた[1]。IoT アバタは、既存のロボットアバタのように、比較的高価なロボットを様々な場所に配置するのではなく、その場に存在するモノをアバタとして遠隔ユーザが操作する。一方、近年では、生成 AI の活用が進展し、IoT アバタに言語生成 AI を組み込むことで、遠隔ユーザの介在なしに現場の人と自律的にコミュニケーションする IoT アバタの実現も可能であり、その開発を行ってきた[2]。

本研究では、複数の AIoT アバタを活用した情報提供システムを提案する。本システムでは、実空間の複数箇所に配置されたモノを AIoT アバタ化し、現場の利用者に対して情報提供およびインタラクティブなコミュニケーションを実現することを目的とする。

### 2. AIoT アバタによる情報提供システム

#### 2.1 システム設計

AIoT アバタは、Raspberry Pi 4 Model B、小型 Bluetooth スピーカー(Kreafunk aGO mini)、広角魚眼カメラ(VR 220 Camera)、小型 USB マイクから構成される AIoT アバタデバイス(図 1)を様々なモノに取り付けることで実現される。生成 AI の実装については、OpenAI Realtime Console と Realtime API を用いて、リアルタイムで応答する Web アプリケーションを構築し、自律的な対話を可能とした。また、プロトタイプとして TensorFlow 及び MoveNet により、AIoT アバタのカメラから人を検出する機能を実装し(図 2)、人を検出した際に AIoT アバタ側から話しかけるシステムについても開発した。



図 1 AIoT アバタデバイス



図 2 人検出の様子

†1 YUKI KIDA, Keio University  
†1 HIDEKI MUTO, Keio University  
†1 TOSHIHIKO CHIBA, Keio University  
†1 TETSURO OGII, Keio University

#### 2.2 複数の AIoT アバタの活用

本システムでは、クリスマスツリー、消毒液のボトル、ホワイトボードに AIoT アバタデバイスを取り付け、AIoT アバタとした(図 3)。AIoT アバタには、それぞれそのモノとして会話し情報提供を行うためのプロンプトを与えた。



図 3 AIoT アバタ

具体的な活用としては、施設内の複数箇所に配置された AIoT アバタによる目的地までの案内・誘導が挙げられる。例えば、ドア入口付近の AIoT アバタがユーザへの案内を開始し、通路の途中に設置された別の AIoT アバタが引きつづき目的地へ案内・誘導を行う(図 4)。



図 4 複数の AIoT アバタの活用

### 3. おわりに

本システムは、将来的に商業施設や公共空間における新たな情報提供手段としての応用が期待される。一方で、本システムの有効性については、実証実験を通じた検証が必要であり、今後の展望として挙げられる。

#### 謝辞

本研究は公益財団法人 JKA の機械振興補助事業(2024M-549)、東京都の先端技術を活用したバリアフリー観光推進事業の補助を受けて実施しました。

#### 参考文献

- [1] Kida, Y., Chiba, T., Ogi, T. (2025). IoT avatar: various objects in real space are anthropomorphised as avatars. *International Journal of Web and Grid Services*, 21(1), 42-57.
- [2] 木田勇輝、武藤英樹、千葉俊彦、小木哲朗 (2025). IoT アバタにおける遠隔ユーザ操作及び生成 AI によるコミュニケーションの比較評価、日本バーチャルリアリティ学会 第30回日本バーチャルリアリティ学会大会論文集、2B2-10.