

博士論文(要約)

Real World 認知科学の実現に向けた
オンライン実験プラットフォーム GO-E-MON の開発

谷澤 智史

目次

第1章 序論	1
1. 研究の背景	1
1.1 Real World 認知科学	1
1.2 Real World 認知科学とパーソナルデータ	2
1.3 Real World 認知科学とオンライン実験環境	4
1.4 Real World 認知科学と学習科学	6
1.5 Real World 認知科学とウェアラブルセンサー技術	7
2. 研究の目的	9
3. 本稿の構成	11
第2章 オンライン実験プラットフォーム GO-E-MON の開発	12
1. GO-E-MON の概要	12
1.1 分散型 PDS による安全性とその欠点	12
1.2 自由度の高い実験タスクの実現と簡便性	12
1.3 外部システムとの連携と安全性	14
1.4 GO-E-MON のアプローチ	15
2. GO-E-MON の設計	16
2.1 GO-E-MON のアーキテクチャ	16
2.2 GO-E-MON のモジュール構成	17
2.3 GO-E-MON の利用者管理	21
2.4 GO-E-MON のタスク定義	22
3. GO-E-MON の実装	28
4. まとめ	29
第3章 実験プラットフォーム GO-E-MON の評価	30
1. 評価実験の概要	30
2. 実験 1: 大学生による PBL 型授業での使用	32
2.1 実験の概要	32
2.2 実験方法	33
2.3 実験結果	38
2.4 考察	44
2.5 結論	45

3. 実験 2: 講義動画に対する学習動態の分析	46
3.1 実験の概要	46
3.2 実験方法	47
3.3 実験結果	49
3.4 考察	55
3.5 結論	56
4. 実験 3-1, 3-2: 英語学習コンテンツを用いた長期学習における学習動態の分析	58
4.1 実験の概要	58
4.2 実験方法: 実験 3-1	60
4.3 実験方法: 実験 3-2	68
4.4 実験結果	71
4.5 考察	78
4.6 結論	80
5. まとめ	81
第 4 章 考察	83
1. Real World 認知科学実験への障壁	83
2. 実験結果データ管理における安全性のさらなる強化	85
3. 実験用スクリプトの記述の簡便性と安全性	87
4. 実験協力者自身の同意によるビッグデータ形成に向けて	88
第 5 章 まとめ	91
1. 結論	91
2. 今後の課題	93
謝辞	95
参考文献	97

図目次

図 1 GO-E-MON のアーキテクチャ	16
図 2 GO-E-MON のモジュール構成	18
図 3 TaskEditor によるタスクの定義	19
図 4 Personary によるログの確認	20
図 5 LTI tool provider との連携フロー	24
図 6 外部システムとの連携フロー	26
図 7 動画視聴タスクのユーザーインタフェースとログ形式	33
図 8 数字記憶タスクのユーザーインタフェースとログ形式	35
図 9 サンプルタスクのデータフロー	36
図 10 解析用 Jupyter Notebook の画面例	37
図 11 動画視聴行動の収集例	38
図 12 講義後アンケートの集計結果	42
図 13 タスクの流れと用意した講義動画	47
図 14 分割型と通し型における点数変化	50
図 15 プレテストの点数帯による点数増加傾向の違い	51
図 16 プレテストの点数帯と平均再生速度, ザッピング回数の関係	52
図 17 プレテスト点数帯と参照チャプター数の関係	54
図 18 実験の流れ及びアプリケーションの構成	61
図 19 マンガリッシュアプリ	62
図 20 英単語アプリの画面イメージ	64
図 21 自由英作文アプリ	65
図 22 実験の流れとアプリケーションの構成	68
図 23 Garmin のデータ収集の流れ	69
図 24 LINE による実施リマインドの流れ	70
図 25 英語に対する自己評価と実施割合	72
図 26 GO-E-MON のリソース取得(GET)所要時間	73
図 27 実験期間中のログ出力量とログ遅延の変化	73
図 28 随伴的な提示とランダム提示における学習量の違い	74
図 29 随伴的な出題とランダム出題による学習継続率	75
図 30 TOEIC 事前点数と睡眠動態の関係	76
図 31 睡眠動態・学習動態と TOEIC 点数変化	77

表目次

表 1 既存のツールと GO-E-MON	15
表 2 GO-E-MON の構成技術	28
表 3 評価実験と評価対象	30
表 4 学生グループによる発表と使用したデータの種類	40
表 5 講義後のアンケート	40
表 6 動画のトピック構成	48
表 7 プレテスト・ポストテスト	66
表 8 アプリ別利用者数	71
表 9 生理指標収集の利用状況	76

特許出願予定の発明を含んでいるため、
第3章4.を削除し、関連する箇所を墨消しした
版です。

第1章 序論

1. 研究の背景

1.1 Real World 認知科学

認知科学は、情報の概念と情報科学の方法論を基礎として発展してきた知的営みである(安西, 2011). 主に実験室実験という形で多くの実験がなされ、人間の認知・行動に関わる全ての機能を情報の処理とみなすことで、多くの実りある成果を残してきた. しかし、これらのデータは実験室という特殊な場での、ある時点での断片的な記録という制約がある. このような特殊な場、特殊な刺激などの制約からの脱却として、開らが提案する Real World 認知科学を指向する必要があると考える(開・今井・嶋田, 2019). Real World 認知科学では、インタラクション研究を重視する. 特に、「二人称的研究パラダイム」として、人間対人間、および機械システム対人間のインタラクション研究において、応答の仕方を実験者が制御可能な実験パラダイムの必要性を強調している.

インタラクション研究の例として、山本らの「ダブル TV パラダイム」を用いた実験が挙げられる. この実験では、2 揃いのカメラとディスプレイを用いて、24 か月児とその母親がカメラとディスプレイ越しにコミュニケーションを行う. この実験では、玩具のデモンストレーター(母親)と被教示者(幼児)の間に、わずか 1 秒のコミュニケーション遅延が挿入されただけで、幼児の模倣学習に影響があることが示された(Yamamoto et al., 2019). この知見は、人間対人間の相互作用の場に対して、遅延を挿入するという制御を加えることで明らかになったものである. Real World 認知科学では、このような二人称的研究を長期にわたって実施し、継続的にそのデータを得ることが重要であるとしている.

COVID-19 禍において対面による実験が困難な時期が続いた一方で、パンデミックは会議や授業をオンライン化せざるを得ない状況を作り出し、タブレット等のデジタル端末やウェアラブルセンサーの普及を促進した. これらの機器の普及により、オンラインでの長期・継続的な実験は実現しやすくなった. 実験者がオンライン実験を実施する場合、質問紙やタスクをオンラインで作成、実施できるオンライン実験サービスを利用することができる. しかし、オンライン実験サービスがサポートしていないような、独自の実験デザインに基づく実験を行う場合は、実験者が自身でオンライン実験システムを構築する必要がある. また、長期にわたる継続的なデータは、オンライン実験サービスに置くには安全性に不安が残る. Real World 認知科学を指向したインタラクション研究を行える実験デザインの柔軟性を持ち、かつプラットフォームで得られる実験データを安全に統合可能な環境を整備することができれば、実世界を指向したインタラクション研究の発展に寄与できる可能性がある.

1.2 Real World 認知科学とパーソナルデータ

Real World=長期・継続的なデータを利用した研究の一例が、Roy らによる研究である(Roy, 2006). 彼らは全方位カメラとマイクを自宅の各部屋に設置し、Roy 自身の子が誕生後すぐから 3 歳になるまでに接触したほぼ全ての言語的データを記録した. この方法を使えば、子どもに与えられた音声刺激、子どもが発した音声データを途切れなく記録することができる. ただしこれらのデータは、実験協力者である子ども及びその家族にとってのプライバシーに関わる情報を多分に含むものであり、この実験は実験者と実験協力者が同一であることで初めて実現できたとも考えられる.

乳幼児の発話記録のようなデータに限らず、位置情報やウェブ検索履歴、睡眠記録などのパーソナルデータは、現代のデータ集約型科学の「油田」であり(Lazer et al., 2009), 現代のオンライン経済の基盤でもある(Schwab et al., 2011). 現在、このような情報は様々なアプリケーションによって収集され、サービスや体験を最適化するために利用されている(Almuhimedi et al., 2015). パーソナルデータの利用は一次利用と二次利用の 2 種類に大別することができる. 一次利用とは、パーソナルデータを、その本人に対するサービスの最適化に用いることである. 一次利用においてそのメリットを享受するのはデータ所有者本人である. 二次利用とは、パーソナルデータを用いて研究開発等を行うことで、有益な知見が得られるといった社会全体がメリットを得ることができる利用形態を指す. 現在は、パーソナルデータの一次利用および二次利用は、Google や Apple 等のビッグテックと呼ばれる、多角的にデジタルサービスを提供する企業を中心に実現されている(Birch, Cochrane, & Ward, 2021). 異なるサービスや企業によって収集されたパーソナルデータは、他の企業が利用してデータ所有者に対する先進的なサービスを実現したり(一次利用)、他の研究者が利用して革新的な研究を実施したり(二次利用)することは困難である. その理由の一つとして、これらの情報を扱うには個人のプライバシーや法的側面への配慮が必要であることが挙げられる(Eagle, Macy, & Claxton, 2010; de Montjoye et al., 2013; Narayanan & Shmatikov, 2008).

これら一次利用と二次利用をどのように実現すべきか、医療における患者の概要、電子処方箋、検査結果等の健康データにおいては、EU では EHDS(the European Health Data Space)が提案されている(板倉, 2023). EHDS では研究者、イノベータ、政策立案者が電子健康データを利用できるようにすること、すなわち二次利用を目的の一つとする. EHDS では、研究者等のデータ利用者と医療機関等のデータ保有者の間に「健康データアクセス機関」と呼ばれる機関を置くことで、健康データの二次利用を実現する. データの二次利用を希望する者は健康データアクセス機関に対してデータアクセス申請を行うことができる. 健康データアクセス機関はその申請を審査し、データ所有者への不利益のおそれがなく、これを許可となった場合は、データ保有者に対してデータの開示を依頼する. データ保有者からデータ利用者に対するデータの開示を、提供義務として規定することで、データの二次利用の実現を図っている. 本論文が対象とする認知科学等における実験データについては、特に実験データを用いて実現しうる一次利用的なサービスの明確なイメージはまだない. そのため、医療のような二次利用のための体制が社会的に整備されていくのにはさらに時間がかかると考えられる.

これらのパーソナルデータの取り扱いに関する問題を解決する方法として、分散型パーソナルデータストア(PDS)が提案されている. 分散型 PDS では、利用者のデータをクラウドサービス等が集中管理

する方式とは異なり、各利用者個人が自分のデータを自身で管理する。Pentland らによる openPDS / SafeAnswers (de Montjoye et al., 2014) は、利用者のデータを個人の端末外に持ち出すことなく、そのデータを活用したサービスを実現できるよう設計された PDS である。利用者は自身の位置情報、検索履歴、通話履歴などの個人データを使ったサービスに参加する場合、サービス提供者が提供する小さなコード(個人データからある計算を行い、結果のみをサービス提供者に返すコード)を、自身の端末内で動作させることができる。このコードによって得られた計算結果のみがサービス提供事業者に送信され、それに基づくサービスが利用者に対して提供される。利用者のデータそのものを端末外に持ち出さなくてもよいデータ処理方式を提供することで、個人データの安全性を実現している。

もう一つの例は、橋田による Personary(Hasida, 2014)である。Personary は、利用者のみが知るパスワードで保護された秘密鍵により暗号化したパーソナルデータを、パブリッククラウドが提供するストレージ上に保存する。利用者しか復号できない方式で暗号化しておくことで、パブリッククラウドによる可用性の恩恵を受けながら、パブリッククラウド組織内の悪意のある攻撃者や漏洩によるデータ窃取を防ぎ、安全性を実現する。Personary は、データをチャンネルという単位で管理する。チャンネルにはタイムスタンプが付与されたアイテムを複数格納することができる。利用者のチャンネルは他の利用者に開示することができる。他の利用者に対してチャンネルを開示する場合、自身と開示先の利用者のみが復号できるようチャンネルを暗号化することで、開示先の利用者もこのデータを参照することができる。いずれの PDS 技術も、利用者は自分のデータを、自身の端末、あるいは自身が所有するストレージ内に閉じて管理をすることができ、どのデータを誰に共有するかは全て利用者自身が決定できる。そのため、利用者は自身の意思で、サービス等に自身のデータを提供することで、自身に最適化されたサービスを受ける(一次利用する)ことが可能である。

実験データの管理に分散型 PDS を用いて、研究への二次利用を容易にすることで、行動科学分野の研究手順の健全性を向上させることも期待できる。例えば、研究結果の「再現性の危機」という問題は、あらゆる学問分野に存在し(Baker, 2016)、医学や心理学の分野では特に深刻だと考えられている。ある調査報告(Open Science Collaboration, 2015)によれば、トップジャーナルに掲載された心理学論文の再現性はせいぜい 50% である。生物や人間を対象とした研究の結果には多くの要因が影響し、特に社会科学の分野では、1 回の研究で必要な条件を完全に制御することは不可能である(Camerer et al., 2018)。研究デザインは実験者の期待に強く依存し、多くの人が関心を持つテーマについて、より厳密な検証に耐える理論と同じくらい頻繁に、新規の(確実ではない)理論が引用される(Yang, Youyou, & Uzzi, 2020)。

このような信頼を揺るがす事態に対処するための主な方法として、以下のアプローチが挙げられる。

- (1) 仮説やデータ解析の計画を事前に開示する(Center for Open Science, 2013)
- (2) 第三者による追試に耐えるように生データを公開する
- (3) 取得データをプールした大規模データに基づくオープンデータ科学に移行する

上記(1)の実施については、最終的に研究者の倫理観に依存するという問題点が指摘されている(Ioannidis, 2016)。また、新しいデータ取得・解析技術がかつてないスピードで開発される中、研究設計

時に既に検証された仮説や解析技術の使用を制限することは、その分野の競争力を低下させ、そのような努力を必要としない商業研究・工学研究に、基礎研究が遅れを取ることに繋がる。

また(2)のデータ開示の要請はもっともであるが、医療情報、社会的差別、個人特定につながる情報など、開示を前提とすると得られないデータも多い。また、先行研究の再現性を検証するためにメタアナリシスを行ったとしても、単純な統合では偽陰性になりやすいことも指摘されている (Doleman, Williams, & Lund, 2019; Hajek et al., 2013)。

大規模データに基づく研究アプローチ(3)に関しては、生物学においては、全ゲノム、発現、脳機能測定データの大規模な蓄積をもとに、バイオインフォマティクスや機械学習を用いて仮説を精緻化することが主流になっている。しかし、心理学や教育学では、データの取得方法が比較的単純であり、人がデータを解釈することも難しくないため、逆にプライバシーの問題が生じ、(3)に対する議論が避けられる傾向にある。実験協力者が、自身が参加した実験に関するデータを、自身の PDS に所有することができれば、実験協力者は自分の意思で、他の実験者に自分のデータを共有できるようになる。このように分散型 PDS を通して実験データ流通が可能な状況を作り出すことができれば、長期にわたる実験 (Trampe, Quoidbach, & Taquet, 2015; Fett et al., 2020; Cragg et al., 2011) で得られた貴重なビッグデータを、本人の同意のもとに仮想的な実験データプールとして統合できる。つまり、心理学や教育学のような領域でも大規模データに基づくデータ科学への移行が期待できる。

1.3 Real World 認知科学とオンライン実験環境

分散型 PDS を実験に利用するためには、利用者の実験課題に対する行動データや、センサーによる生理指標を収集する機能が必要である。openPDS / SafeAnswers は PDS へのデータ入力機能は定義しない。実験に使用するためには、スマートフォンあるいはデスクトップアプリケーションとして、実験課題を利用者に提示しその結果を PDS に保存する処理や、センサーデータを取得して PDS に保存する処理を実験者が実装する必要がある。実験者は独自の実験課題を実装することが可能であるが、アプリケーション開発と配信に取り組まなければならない。

Personary はスマートフォンと PC に対応したスタンドアロンアプリケーションの形で提供される。Personary アプリは、歩数等のスマートフォンのヘルスデータを PDS に保存する機能と、任意の入力欄を定義可能なフォームによるアイテムの編集機能を提供する。Personary アプリは一般に配布されており、アプリケーション開発の手間はかからない。フォームの内容は実験者によって定義可能であるため、質問紙形式の実験が可能である。しかし、心理学実験で使われる実験課題のような、反応時間の取得を行うタスクなどを記述することは実験者だけでは難しく、Personary アプリでの実装が必要となる。

実験者が独自に考案した実験タスクをオンライン実験として提供可能にするためには、Web ブラウザの活用が考えられる。Web ブラウザが高い表現能力を持つことに加え、特別なアプリケーションのインストール作業なしに、標準の Web ブラウザのみでオンライン実験に参加できる利点もある。そこで本研究では、分散型 PDS 技術を用いてパーソナルデータの保護を実現しつつ、Web ブラウザにより利用可能なオンライン実験プラットフォームについて検討する。

オンライン実験, 特に World Wide Web を用いたオンライン実験の歴史は古く, 1999 年の Johnson-Laird らの実験 (Johnson-Laird & Savary, 1999) に始まる. 最近ではオープンソースのオンライン実験フレームワークやソフトウェア (Peirce et al., 2019; Stegman, Crawford, & Gray, 2018) が広く公開されていることで, サーバーやスクリプトに関する一定の知識さえあれば, オンライン実験を容易に行うことが可能になった. Web ブラウザで刺激提示を行う場合, 非常に短い視覚刺激の提示(目安として 20ms 以下)を除いては概ねタイミングの遅れなどに実験上の問題がないことが報告されている (Stewart, Chandler, & Paolacci, 2017). オンライン実験の利点として, 実験協力者と実験者の間の物理的な距離に制約されることなく参加者募集ができるため, 従来の実験室実験で指摘されていた実験協力者の偏りを解決できる可能性 (Masuda, Batgerel, & Senzaki, 2020) があるが, 欠点も指摘されている. 特にオンライン実験では, 実験協力者は実験者と対面していないために注意散漫となることや, 実験室のような専用の空間ではないために実験実施時に邪魔が入りうることで, データの質が低下することが指摘されている (Yetano & Royo, 2017). しかし, このような実験参加のプレッシャーがない環境は, 実験協力者の自主的な行動の変化を記録したいと考える場合には利点ともなりうる. オンライン実験は実験室に対して利点・欠点の両面があり, その性質を理解して注意深く実験を設計することが重要である (Gagné & Franzen, 2023).

Web ブラウザにより実施可能なオンライン実験を実装するためのライブラリとして jsPsych (de Leeuw, 2015) がある. 質問紙を実現するフォーム型インタフェース以外にも, 心理学, 認知科学でよく知られたタスクがさまざまな研究者により jsPsych プラグインとして公開されている (Kuroki, 2021). jsPsych は非常に自由度の高いフレームワークであり, 実験者のスクリプト作成の負担を軽減することができる. しかし, オンライン実験システムを構築するためには, プログラムの作成以外にも, サーバーへのプログラムの配置, ソフトウェアの脆弱性への対応など, 実験者にとって負担となる作業が必要である.

実験者は自前のサーバを用意する代わりに, オンライン実験のためのサービスを利用することでこれらの負担を軽減することができる. 例えば, Amazon Mechanical Turk (Amazon Mechanical Turk, 2005) は, コンピュータでは困難な作業を人々に依頼するために開発されたサービスである. Amazon Mechanical Turk は Web ブラウザを用いて利用でき, アンケートや認知能力を試す作業をスクリプトとして作成し, タスクとして登録すると, Amazon Mechanical Turk に登録した利用者にそのタスクを実施してもらうことでその結果を取得することができる. PsyToolkit (Stoet, 2017) は認知心理学実験に特化したオンライン実験ツールで, 質問紙をはじめ, N-back 課題など認知心理学実験でよく知られた課題をサービスとして利用することができる. オンライン実験にこれらのサービスを活用することで, 実験者に要求される技術的知識や運用の負担を軽減することができるが, 一般的にこれらのサービスを構成するサーバーやデータベース等はサービス提供事業者により管理され, 実験データはサーバー・データベース内で集中的に管理されている. このような実験データの集中管理には, 攻撃による実験データの窃取といったセキュリティ上のリスクがつきまとう.

Web ブラウザを用いたオンライン実験環境の欠点として, 実験協力者側に Web ブラウザでタスクの URL を開いてもらわなければならない点が挙げられる. 長期にわたって継続的に実験協力者からデータを得る場合, プッシュ通知などを使って実験協力者に実施を促すことは必要不可欠である. プッシュ

通知は、Ecological Momentary Assessment(EMA)(Shiffman, Stone, & Hufford, 2008)等の経験サンプリングによる実験では欠かせない。他にも、SEMA3(Koval et al., 2019)のような専用アプリケーションを用意することでプッシュ通知を実現した例や、SMSのようなメッセージサービスを用いて Web アプリケーションへの誘導を行う SurveySignal(Hofmann & Patel, 2015)の例がある。Web ブラウザでのプッシュ通知の実装は Push API(World Wide Web Consortium, 2022)があるものの、現在のところ iOS デバイスでは試験的にサポートされている程度であり(Apple Inc., 2023)、正式にサポートされたとしてもスパムメッセージなど利用者が意図しない行為を防止するため、限定的な運用になる可能性が考えられる。それ以外に、実験協力者の機器に常駐して、通知を定期的に出すようなアプリケーションも考えられる。しかし、スマートフォンやタブレットのオペレーティングシステムは機器が持つバッテリーを有効活用するために、できる限り使用していないアプリケーションを停止することでバッテリーを節約しようとする(Chen et al., 2015)ため、機器内に常駐し続けるようなアプリケーションの実現も困難である。このような問題に対応するため、オンライン実験プラットフォームは外部のメッセージングサービスと連携し、プッシュ通知の発行を容易にできることが望ましい。

1.4 Real World 認知科学と学習科学

Real World 認知科学の適用分野として考えられるものの 1 つに、学習科学の分野がある。2019 年からの COVID-19 禍により日本では教育環境のデジタル化が急激に進み(文部科学省初等中等教育局, 2021), Educational Technology (Edtech) (Weller, 2018)はますます注目されている。しかしながら、教育に関して「いつ」・「どのように」・「何を」教えるべきかに関して、学習者側のエビデンスは十分ではなく、デジタル技術によりどのように教育サービスを効率化し、高度な教育を実現していくかには多くの課題が存在する。

このような課題に対するアプローチの一つとして、学習システムの中で収集されるデータを用いた Learning Analytics がある。Learning Analytics は学習内容やその環境の理解、最適化を目的として、学習者のデータやその状況を計測、収集、分析、利用(Siemens & Long, 2011)するものである。一方でデジタル技術を用いた教育サービス、特にオンライン学習に関する研究に対しては、1990 年代ごろから、実験的介入の欠如が問題点として指摘されている(Phipps & Merisotis, 1999)。観察研究のような介入を伴わない研究は、観測した変数間の相関を明らかにすることは可能であるが、特に学習研究においては学習者や学習環境、教材などの多くの変数を考慮する必要があるため、変数間の因果関係にまで言及することは非常に難しい。現在でも介入を伴う研究が徐々に増えているものの、介入の対象は学習ダッシュボードと呼ばれる学習教材への誘導を行う部分に偏っていることも指摘されており、教材そのものや教師への介入は現在でも非常に少ないことが指摘されている(Tepgeç & Ifenthaler, 2022)。

Real World 認知科学におけるインタラクション研究の観点では、デジタル化された学習教材が学習者に対してどのように接するべきかは、機械システム対人間の相互作用の問題と捉えることができる。学習者に対して、その関心や習熟度に応じて最適な指導やフィードバックを提供する学習システムとして Intelligent Tutoring System(ITS)がある。ITS として有名なシステムに、Corbett と Anderson による LISP Tutor がある。LISP Tutor は、学生の演習中に間違いを特定し、有効なフィードバックを提供することが

できた(Corbett & Anderson, 1992). 学習者の状況に応じて, どのようなフィードバックを提供することが有効なのかを知る上で, インタクション研究は不可欠である. しかし, 一般的な学習アプリケーションは応答の仕方を自由に変更することはできず, 実験への適用は難しい. 応答の仕方を制御可能なアプリケーションを実験のためだけに実装することは, 開発や教材制作のコストがかかりすぎる.

学習アプリケーションをどのように用意するかは, Learning Tools Interoperability (LTI) (1EdTech Consortium, 2019)と呼ばれる規格を活用することで解決できる可能性がある. LTI は, Moodle(Moodle HQ, 2002)や Sakai(Sakai Project, 2005)のような Learning Management System(LMS)と, デジタル学習教材の間で, 利用者の情報や教材による学習結果をやりとりするための標準仕様である. LTI Tool Provider 仕様に基づいて実装されたデジタル学習教材は, LMS 内に容易に組み込むことができ, パラメータを与えることでカスタマイズができる. LTI Tool Provider として実装されたデジタル学習教材には文部科学省の MEXBIT(文部科学省総合教育政策局, 2021)や京都大学の Bookroll (Kannan et al., 2022)がある. LTI Tool Provider と容易に連携可能なオンライン実験プラットフォームを実現できれば, 従来の認知科学, 心理学においてよく知られたタスクだけでなく, 日常的に使用される学習教材を使った時の反応もデータとして統合できる可能性がある.

1.5 Real World 認知科学とウェアラブルセンサー技術

インタクション中の実験協力者の反応を記録するためには, 脳波計や心拍計など, 協力者の生理的な指標を記録できるセンサーが欠かせない. しかしながら, 一般に実験室実験で用いられる脳波計や心拍計などの各種センサーは多くが非常に高額であり, 測定のための装着ノウハウなど利用においても専門知識を要し, 日常的に用いることは困難である. その代替として, 現在市販されている民生用のウェアラブルセンサーを利用することが考えられる. 例えば, 腕時計型のウェアラブルセンサーは, 睡眠や脈波といった生理指標を得ることができ, 実験に利用可能であることが示されている(Evenson, Goto, & Furberg, 2015). また, 携帯型の脳波計を用いた実験も行われている(Poulsen et al., 2017; Sawangjai et al., 2020). これらのセンサーは装着の負担が少ないため, 長期的, 継続的なデータ取得の可能性を秘めている.

学習研究において生理指標を取得, 活用した例としては, スマートフォンから得られる睡眠記録などのデータから学生の成績平均値(GPA)を推定する手法の提案(Wang et al., 2015)や, ウェアラブルセンサーで取得した睡眠記録と, アンケートで取得した知覚ストレス尺度(PSS)から学生の GPA を推定する方法の提案(Sano et al., 2015)がある. これらの実験は GPA のような総合的な点数のみを対象としており, なぜ睡眠記録から成績が推定できるのか, 記憶や認知能力などの認知科学的知見による説明には至っていない. なぜ成績が向上したのかを説明するためには, その結果に至るまでの学習者と教材の関わりを, 例えばフラッシュカードに対する学習動態の変化(Streeter, 2015)などの個別のタスクに関するログとして収集, 分析する必要がある.

認知科学における知見と学習研究を結びつける例として, 「いつ暗記(学習)をするのがよいか」に関する研究が挙げられる. 記憶は単語の暗記など学習における重要な要素であり, 特に睡眠が記憶の定着に重要な役割を果たすことが示されている(Rasch & Born, 2013). 例えば Gais らは, 睡眠の 10 時間

前に単語を暗記するのに比べ、睡眠の 3 時間前に単語を暗記すると、より良い再生成績が得られることを示した(Gais, Lucas, & Born, 2006). しかし、この実験は、実験室での記憶課題と睡眠日誌を用いて実施されたものであり、実験に使用した課題には被験者が触れたことがない言語のものが用いられていた。

現実の学習の場においても同じ傾向が得られるか、つまり入眠前に学習する方が有利かどうかを検討した実験がアクティビティトラッカーとスマートフォンアプリケーションを用いた実験である(Yazawa, Yoshimoto, & Hiraki, 2018). この実験では、アクティビティトラッカーによる睡眠記録と、時間帯ごとに出現単語が偏るように設定された英単語学習アプリケーションを用いて、それぞれの英単語に関して、入眠の何時間前に学習したかと英単語の再認成績の関係が分析された。学習者全体としては入眠前の学習が有利である、すなわち先行研究と同様の傾向が得られたものの、個人差も大きく、逆の傾向を持つ学習者も確認できた。このような実験で得られた、それぞれの実験協力者に関する学習ログや生理指標を分散型 PDS に格納し、二次利用すなわち縦断的な分析ができるようになれば、このような個人差がなぜ生じたのかに関してさらに分析をすることができる。例えば、別の調査で得られた、時間帯ごとの環境音に関するデータがあれば、ある被験者に関しては、「入眠前にいる環境は騒音など学習に適さない環境であるため、入眠前の学習が不利だった」といった説明ができるかもしれない。このような知見を積み重ねることで、入眠に対する最適な学習タイミングを決定する技術を確立できれば、多くの学習者に対して適用可能な学習支援技術を実現でき、社会に対してメリットをもたらすことができる。

ウェアラブルセンサーで得られたデータを収集するためには、連携するシステム開発が必要となる場合が多い。腕時計型のウェアラブルセンサーや、携帯型の脳波計は、スマートフォンやコンピュータと Bluetooth を通じて通信するのが一般的である。しかし民生用のウェアラブルセンサーは、Bluetooth 上の通信プロトコルは公開せず、自社の連携アプリケーションのみが独占的に機器に接続できるようにしている場合が多い。これらのアプリケーションは機器から得たデータを、一度事業者のサーバーに集約した上で利用可能とする(Garmin Ltd., 2021; Fitbit LLC, 2023). これらのウェアラブルセンサー事業者のサーバーからデータを取得するためには、OAuth 2.0(Internet Engineering Task Force, 2012)などのプロトコルを通じてウェアラブルセンサーのアカウント所有者の許諾を得る必要がある場合が多いが、このような機能の実装は OAuth 2.0 にしたがって振る舞うサーバーの準備が必須である。Apple 社が提供する HealthKit API(Apple Inc., 2014)はこれらのウェアラブルセンサーから得られた情報を集約し、iOS 上で動作するアプリケーションに提供する役割を持つが、この API へとアクセスするためには iOS アプリケーションとして実装する必要がある。

これらの連携機能の実現は、ネットワークプロトコル等に関する技術的な知識が必要となる。そこでウェアラブルセンサーとの連携機能を、オンライン実験プラットフォームがサービスとして提供できれば、より多くの実験者が Real World 認知科学を指向した実験を実施できるようになる可能性がある。

2. 研究の目的

本研究では、Real World 認知科学を指向したインタラクション研究を広く実施できる、オンライン実験プラットフォーム **GO-E-MON (Good Experiment for Mankind Online)** を提案する。GO-E-MON は以下の 5 つの特性を備えるものとする。

- (1) 実験課題の柔軟性: 実験者は自身で考案した実験タスクを定義することができ、タスクに関わる操作データやウェアラブルセンサーによる生理指標を収集できる。
- (2) データの安全性: 実験協力者のパーソナルデータである、長期にわたる行動データや生理指標を安全に管理できる。
- (3) 実験環境の簡便性: 実験者はシステム開発やサーバー運用の負担を負わずに実験を実現できる。
- (4) 実験環境の可用性: 長期的・継続的な実験のため、多くの利用者からのアクセスや、機器故障などのトラブルがあっても実験環境を提供し続けられる。
- (5) データの流通性: 実験協力者の同意のもと、実験データを別の実験者が統合することができ、多くの協力者の長期にわたるデータを縦断的に分析することができる。

GO-E-MON は、実験者に対して Web ブラウザで利用可能な実験環境を提供し、かつ、データの保存には分散型 PDS を組み合わせることで、これらの特性を実現する。

GO-E-MON は、Web アプリケーションとして提供する。ブラウザでのタスク実行機能に加え、ウェアラブルセンサーや学習アプリケーションとの外部システム連携機能を提供する。Web ブラウザの表現能力と外部システム連携機能により、実験者による独自の実験デザインの実現を支援する(実験デザインの柔軟性)。GO-E-MON は、Web ブラウザでのタスク実行に必要なファイルを高速に配信し、外部システムとの連携プロトコルを処理するサーバー群を提供する。これらの機能は実験者に対してサービスとして提供され、サーバー群の構築や管理を実験者は意識する必要がない(実験環境の簡便性)。また、これらのサーバー群は、多くの実験協力者がいつでも利用可能な状態を維持する(実験環境の可用性)。

実験により得られたデータは、GO-E-MON を構成するサーバー群に長期に保存することはない。実験結果データは取得後速やかに実験者および協力者の PDS に格納し、GO-E-MON サーバー群からは削除することで、データの分散管理を実現する(データの安全性)。また、実験協力者は自身の責任において、自身が参加した実験データを所有、管理することができ、自身の自由意志に基づいて別の実験者にデータを提供することができる(データの流通性)。

GO-E-MON は Real World 認知科学がカバーしうる領域全て、例えば発達研究などの様々な領域への貢献を目的としたものであるが、本研究では GO-E-MON が Real World 認知科学を指向した実験に利用可能であることを示すための事例として、4 つの実験について述べる。高校生、大学生を実験協力者とした実験の事例から、GO-E-MON が 5 つの特性を備えているかを検討し、そこで得られた課題について示す。

まず、実験者と実験協力者が GO-E-MON と Personary の操作を理解し、使用することができるかを検証することを目的とし、実験を行った(実験 1)。実験 1 では、講義に参加した大学生 19 名を実験協力者として、GO-E-MON と Personary を参加者が利用でき、データの安全性と流通性が実現できるかを調査する。GO-E-MON 上で、実験協力者にタスクを実施してもらい、得られたデータを実験者および協力者が自身の PDS で分散管理できることと、実験協力者によるデータの開示操作が可能かどうかの調査を行なう。

GO-E-MON で可能なことと課題を示すための事例として、実際に GO-E-MON を用いて行った 3 つの学習研究に関する実験について述べる。GO-E-MON が柔軟にタスクを構成できることを示す事例の一つとして、実験 2 を紹介する。実験 2 は、GO-E-MON 開発者とは異なる実験者が、高校生 78 名を実験協力者として、講義動画視聴システムに対する学習者の行動を調査したものである。実験者は講義動画の作り方による学習効果を調査することを目的とし、動画プレイヤーとプレ・ポストテストが可能な実験タスクを実装し、実験を行うことができた。

GO-E-MON がウェアラブルセンサー等の外部システムと連携する柔軟性を持ち、長期間にわたるデータが継続的に収集できる可用性があることを示す事例として、実験 3-1 と実験 3-2 について述べる。実験 3-1 と 3-2 は、学習者の学習行動によって出題内容を変更することの学習継続効果を調査することを目的とし、独自に英語学習アプリケーションを開発し、高校生 1422 名、大学生 20 名を対象に実験を行なったものである。実験 3-1 は外部学習アプリケーションとの連携により複雑な実験タスクを実現し、実験 3-2 はウェアラブルセンサーとメッセージングサービスとの連携を行った。実験 3-1 では多人数の利用でも GO-E-MON は継続的にサービスを提供し続けられることを確認することができた。

これらの事例から、オンライン実験プラットフォーム GO-E-MON が、Real World 認知科学の実現のために必要な特性を備えることを示す。また、実験を通じて得られた、Real World 認知科学を指向した実験を多くの研究者に実施してもらう上での解決すべき課題についても議論する。

3. 本稿の構成

本稿はまず第 2 章として, Real World 認知科学を指向した様々な形態の実験を, 分散型 PDS 技術により安全に実施できるオンライン実験プラットフォーム GO-E-MON の開発について報告する.

次に, 第 3 章として, GO-E-MON がオンライン実験プラットフォームとして利用可能であることを検証するため, 大学生を実験協力者としてのケーススタディを実施したのち, 高校生・大学生の学習場面における実験を実施した. GO-E-MON が Real World 実験環境としての特性を備えることの検証結果と合わせて, これらの実験によって得られた知見についても報告する.

第 4 章としてこれらの実験結果について考察し, GO-E-MON の可能性と課題について述べる. 最後に第 5 章として結論をまとめる.

第2章 オンライン実験プラットフォーム GO-E-MON の開発

1. GO-E-MON の概要

GO-E-MON は Real World 認知科学のコンセプトに基づき、実験協力者対機械システム、あるいは人間対人間の間の相互作用における行動記録や、センサーによる生理指標データを、長期的かつ継続的に取り扱う必要がある。これらのデータは実験協力者にとってパーソナルデータであり、集中管理型の実験プラットフォームでは安全性の面で不安が残る。一方で広く利用可能で、かつサーバー等の準備の負担を実験者に負わせない簡便性を備えようとする、コンテンツの配信機能や、外部サービスとの連携機能を担うサーバー群は、サービスとしてどこかに集中的に管理せざるを得ない。本セクションではオンライン実験の実現方式の利点と欠点をまとめ、GO-E-MON のアプローチについて整理する。

1.1 分散型 PDS による安全性とその欠点

GO-E-MON は分散型 PDS と連携するオンライン実験サービスとして設計する。openPDS / SafeAnswers はフレームワークであり、スタンドアロンアプリケーションに組み込むことで分散型 PDS としての機能を提供することができる。一方で Personary はスタンドアロンアプリケーションの形で提供され、Personary アプリはスマートフォンのアプリストアから誰でもインストールすることができる。

GO-E-MON は、分散型 PDS として、一般に公開されているため容易にインストールが可能な Personary を採用する。しかし先述の通り、Personary アプリで、実験者が実施可能な実験課題は Personary の入力フォーム仕様に制約される。実験者による独自の実験タスク定義を実現するためには、Web アプリケーション型の実験プラットフォームのように、実験者が独自に画面構成や動的な振る舞いを定義できるような柔軟性が必要である。

比較的複雑な実験タスクを提供し、かつ利用者個人が実験データを管理できるようにした事例として、脳波実験サービスプロバイダである MYND(Hohmann et al., 2020)が挙げられる。民生用の脳波計を使い、MYND 上で実験タスクを実行可能にすることで、個人が自分でタスク中のデータを計測・管理できるようにしている。しかし、MYND は脳波および脳波関連の実験課題に利用が限定されていること、実験者は MYND の運用者のみであり、別の実験者が MYND 上で動作する独自のタスクを作成できないという点で本研究とは異なっている。

1.2 自由度の高い実験タスクの実現と簡便性

実験者に自由度の高いタスク定義を許すためには、Web ブラウザの利用が最適である。Web ブラウザは自由度の高いレイアウトに加え、JavaScript (ECMAScript) (Ecma international, 1997)を用いて複雑な振る舞いを定義することができる。GO-E-MON が実現を目指す Real World 認知科学は長期間、多人数の縦断データを収集することを想定しているが、Web ブラウザはさまざまな機器にインストールされているため、広く多様な利用者に、多くの場面で機能を提供できるという点でも適している。実験タスクの作成機能も Web ブラウザで提供できれば、実験者は特別なツールをインストールすることなく、簡単に実験タスクの作成を行うことができる。

Web ブラウザは、利用者から指示された URL に含まれるドメイン名によりインターネット上のサーバーを特定し、サーバーからページを表現する HTML ファイル等のリソースを取得することで画面表示等を行う。そのため、Web アプリケーションとして実験システムを提供するためには、インターネットに接続されたサーバーを用意する必要がある。GO-E-MON はこのサーバーの管理を引き受け、実験者に対してサービスとしてリソースの配信機能を提供することで簡便性を実現する。サーバーを管理する上で重要な要素が、サーバーが常に利用できる状態になっていること、すなわち可用性の確保である。簡単に利用できたとしても、サーバー群の動作が安定せず、継続的に実験を実施できなければプラットフォームとして成立しない。多くの要求を処理可能なシステムを稼働させ続けることは決して容易ではない。例えばオンライン学習プラットフォームでは 2020 年に発生した教育プラットフォーム Classi の大規模障害(Classi 株式会社, 2020)といったケースもあり、可用性は容易には実現できない要素である。

高い可用性を容易に得られるアプローチの一つとして Amazon Web Services (AWS) (Amazon Web Services, Inc., 2006), Google Cloud Platform(GCP)(Google LLC, 2008)のようなクラウドサービスの活用がある。クラウドサービスを活用して GO-E-MON サーバー群を構成しようとした場合、PaaS(Platform as a Service), または IaaS(Infrastructure as a Service)と呼ばれる 2 種類のサービスの利用が考えられる (Sareen, 2013)。クラウドサービス側に責務を持たせる PaaS 方式であれば、可用性の確保において GO-E-MON が果たすべき責務を小さく抑えることができ、クラウドサービスが持つ可用性の恩恵を得やすい。例えば、GO-E-MON が配信すべきデータをあらかじめすべてクラウドサービスに指示しておけば、クラウドサービスによって適切に多重化された配信機構が準備され、リソースの高速な配信や、機器の故障に対する耐性を得ることができる。しかし一方で、保存されたデータは原則としてクラウドサービスから参照可能な状態になるため、契約に基づいてクラウドサービス事業者によって情報の保護はされるものの、バグや脆弱性などで情報を窃取される危険性が残る。

対して安全性に関して自由度が高いのが、仮想マシンと呼ばれる形態でサーバーを借り出し、サーバー群として利用する IaaS 方式である。この形態では、ディスクや通信路を適切に暗号化することでクラウドサービス事業者がアクセス可能なデータを制限することができる。しかし、仮想マシンやそれらが接続される仮想ネットワークの可用性はさほど高くなく、これらがダウンした際にどのようにサービスを継続するかといった冗長化設計、回復方法の策定などの、可用性確保のための設計の責務は GO-E-MON に委ねられることになる(Mesbahi, Rahmani, & Hosseinzadeh, 2018)。

アプリケーションに関する全てのストレージを Personary により実現することが理想であるが、Personary は Google Drive 等のパブリッククラウドのストレージにデータを保存するため、データの読み込みおよび保存に要する時間はパブリッククラウドのストレージ性能(Akter et al., 2018)に制約される。また、パブリッククラウドの API には、一定時間内にアクセス可能なリクエスト上限数に一定の制約があり、多人数が同時利用するアプリケーションのストレージとしては要件を満たさない。

柔軟な実験デザインを実現するための Web アプリケーション配信機構を、実験者の負担がないようサービスとして提供することを考えた場合、安全性と可用性はトレードオフの関係にある。そこで、アプリケーションが保持すべきデータを (a)高速に読み書きが必要であるが短期的に維持すれば良いデータと (b)低速な書き込みでも問題なく長期的に維持すべきデータ に分け、(a)はクラウドサービスが持つ高

速なデータストアに保存する代わりに保持期間の制約を設けることで流出の危険を軽減し、(b)を Personary に保存することで安全性と高い可用性の両立を目指す。

1.3 外部システムとの連携と安全性

Web アプリケーションの形で提供することで、ブラウザ上での実験タスクの振る舞いは実験者が自由に制御できるようになるが、ブラウザ外のシステムとの連携に対しては強い制限がある。この制限はブラウザが稼働している機器やネットワークを保護する目的で設けられている。そのため、継続的なインタラクション研究が必要とする機能のうち、ウェアラブルセンサーとの連携や、実験協力者への通知のためのメッセージングサービスとの連携、外部の教育アプリケーションとの連携などは、Web ブラウザのみでは実現しづらい。

多くの場合、これらの外部サービスとの連携には、いったん外部サービスの認証あるいは認可画面へと実験協力者を誘導し、その画面上で協力者の許可を得た(権限の委譲を受けた)上で、外部サービスからその権限と紐づいたトークンを受け取ることで連携を実現する。これらの認証、認可依頼、トークン受領処理といった処理の実装には専門的な知識が要求されるが、これらの処理を GO-E-MON がサービスとして提供することで、実験者が外部サービスとの連携を簡単に利用することができるようにする。連携に必要な情報は外部サービスとの連携の際にすぐに読み込み、利用できるよう、クラウドサービスが持つ高速なデータストアに保存する。ただし、これらのトークンが流出するなどして悪用されることができないよう、GO-E-MON はこれらのトークンが長期的にデータストアに残存することがないように設計する。

これらの機能の実現には専用のサーバーを用意する必要があるが、これらを GO-E-MON があらかじめ用意することで、ウェアラブルセンサーからのデータを受信してタスクの動作に反映させたり、EMA (Shiffman, Stone, & Hufford, 2008)のような経験サンプリングのためのメッセージを定期的送信したり、LTI に対応した学習教材アプリケーションを課題として使うといったことが容易に実現できる。

1.4 GO-E-MON のアプローチ

既存ツールに関する利点および欠点を、実験デザインの柔軟性、実験環境の簡便性と可用性、データの安全性、実験データの流通性の観点で整理し、GO-E-MON と比較する(表 1)。

表 1 既存のツールと GO-E-MON

ツール	アプリケーション 提供形態	データ 管理者	柔 軟 性	安 全 性	簡 便 性	可 用 性	流 通 性
JavaScript ライブラリ(jsPsych)	Web アプリケーション	実験者	○	△		△	
オンライン実験環境 (Amazon Mechanical Turk)		サービス提供者, 実験者	○		○	○	
openPDS / SafeAnswers	スタンドアロン (スマートフォン/PC)	実験協力者	○	○		○	○
Personary	アプリケーション			○	○	○	○
GO-E-MON	Web アプリケーション	実験者, 実験協力者	○	○	○	○	○

jsPsych のような JavaScript ライブラリは、実験者によるスクリプト次第で、実験協力者との複雑なインタラクションや、外部システムとの連携が可能である。データの管理に関しては実験者が制御できるが、サーバー等を自身で管理する分、実験者の負担は大きい。可用性をどの程度実現できるかは実験者の技術的な知識や、十分な台数のサーバーを用意できるかといった点に依存する。Amazon Mechanical Turk のようなオンライン環境では、実験者はサーバーの準備をすることなく容易に実験を実施できる。また、可用性もサービスによって確保される。しかし、データはサービス事業者のサーバー上に保持されるため、データ管理の責務をサービス提供事業者に委ねなければならない。実験者が収集した実験データを、その実験外で利用することは研究倫理や許諾取得の点で困難である。

分散型 PDS である openPDS / SafeAnswers や Personary はデータの安全性と流通性を備える一方、openPDS / SafeAnswers はスタンドアロンアプリケーションとして実装しなければならない。実装次第で様々な実験タスクが実現できるが、アプリケーションの開発や配信の知識が必要となるため簡便性に欠ける。Personary はスタンドアロンアプリケーションとして配信されているため利用は容易であるが、実験協力者からのデータ入力は Personary アプリケーションが規定したものに限られるため、実験者が実現可能な実験デザインの柔軟性は Web アプリケーションに比べて低い。

GO-E-MON は、実験者に対して Web ブラウザで利用可能な実験環境を提供し、かつ、データの保存には分散型 PDS、特にアプリとして配布されており、利用が容易な Personary を利用することで、Real World 実験環境が満たすべき特性の両立を目指す。

2. GO-E-MON の設計

2.1 GO-E-MON のアーキテクチャ

GO-E-MON は、Web ベースのオンライン実験プラットフォームに、分散型 PDS への実験データ保存機能を追加する。これにより、実験者による安全な実験データ収集と、従来は不可能であった実験協力者による実験データ管理を実現する。GO-E-MON はこれらの機能を持つ Web アプリケーションサーバーをサービスとして提供する。オンライン実験配信に必要なリソース配信機能や外部サービスとの連携に必要な管理機能は全て GO-E-MON がサービスとして提供することで、実験者は自身でサーバー等を構築することなく、実験用タスクを作成、配信できる。

GO-E-MON 上の実験タスクにより得られた実験結果データは、アプリケーションサーバーに永続的には保存されず、実験者の PDS(Personary)と実験協力者の Personary に送信される(図 1)。

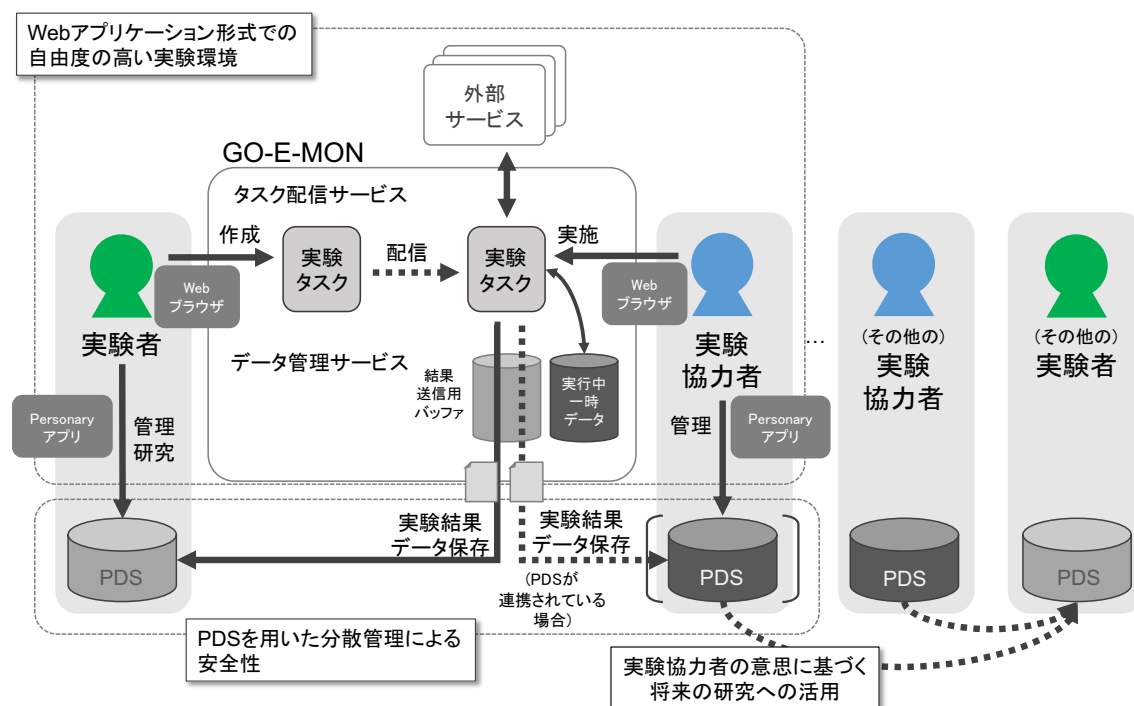


図 1 GO-E-MON のアーキテクチャ

実験者は GO-E-MON を通じて、実験協力者に実験タスクを配信することができる。タスクの実行に必要な、外部システムとの連携用アクセストークン等の情報は、必要な時にすぐにアクセスできるようにするため GO-E-MON サービス内の一時データ用データストアに保存する。この一時データは安全性を考慮して、長期に維持することせず、アクセスがないものを定期的に削除する。長期に維持すべき実験データは実験者と実験協力者が持つ Personary に保存する。GO-E-MON は実験協力者がタスクの実施を完了したのち、それぞれの Personary に実験結果データを追記する。なお、実験協力者が自身の Personary を GO-E-MON に連携するかは実験協力者本人の自由意思に基づくものとする。

Personary への保存後即時にデータをサーバーから削除することで、一元管理されたアプリケーションサーバへの侵入による大規模なデータ漏洩を防ぎ、データの安全性を確保する。

GO-E-MON は、実験者と実験協力者の双方の Personary に実験データを送信する形式とした。GO-E-MON は実験協力者の Personary のみに実験結果データを送信し、実験協力者がそのデータを実験者に開示する形式も考えられたが、全ての実験協力者に自身のデータの管理を強要することになるため、採用しなかった。現在、Personary を利用するためには、パブリッククラウドのストレージサービス、特に Google アカウントにアクセスし、スマートフォンや PC の Personary アプリケーションを使って自分のストレージを作成する必要がある。しかし、実験協力者が Google アカウントを持っているとは限らない(ZDNET, 2020)。また、日本ではスマートフォンや PC の所有率が高いとは言え、実験協力者がアプリケーションに対応したデバイスを所持していない可能性もある(Statistics Bureau, Ministry of Internal Affairs and Communications of Japan, 2020)。そのため、実験協力者によっては Personary の操作を依頼することが困難な場合がある。また、Personary の操作をするのに十分な能力があったとしても、実験結果データを「持ちたくない」と考える実験協力者や、「自身のデータを持つのは面倒だ」と考える実験協力者も一定数いるものと予想される。実験参加の前提として「実験協力者が Personary を操作可能であること」を設定することで、特定の年齢層や背景を持った実験協力者のみがデータを提供する形になり、結果的に、ウェブ実験で指摘される参加者の偏り(日本学術会議, 2020)が、より強くなってしまう可能性が考えられる。そのため、実験協力者による Personary の管理は任意とする。

複数の実験者が得た実験データを実験者間で容易に個人照合、すなわち名寄せできないようにするため、実験タスクと実験協力者ごとに異なる仮名 ID を生成し、その仮名 ID をデータに付加して、実験者の Personary に格納する。データに対して、メールアドレスやユーザ ID などの実験協力者を特定可能な識別子を付加することはない。ユーザ ID ではなく仮名 ID を用いることで、万が一、ある実験者の Personary 内の実験結果データが漏洩したとしても、そのデータが他の実験者のデータと容易に突合されることを防ぐ。

実験協力者の Personary に蓄積された実験データは、実験協力者が自身の意思に基づいて管理・活用することが可能である。Personary にはチャンネルの共有機能があり、実験協力者が別の実験者に、自身の過去の実験結果データを開示したい場合は、チャンネル共有機能によってデータを共有することができる。実験協力者の Personary に蓄積された実験結果データにも、実験者に送信されるものと同様に個人が識別できる ID 等は付与されない。異なる仮名 ID を持った実験結果データであるが、実験協力者自身の個人のチャンネルに集約されているため、このチャンネル内では名寄せされている状態となる。このデータを分析すれば、個人の実験における行動に関して、長期にわたる変化を追うことができる。

2.2 GO-E-MON のモジュール構成

GO-E-MON は、4 つのソフトウェアモジュールにより構成される。オンライン実験スクリプトの編集と管理を行う TaskEditor と、作成したタスクを実行する TaskView、TaskView によって得られた実験データを実験者と実験協力者の Personary に送信する ResultStore、そして外部サービスと接続するためのプロトコルを実装する External Service Connector である(図 2)。

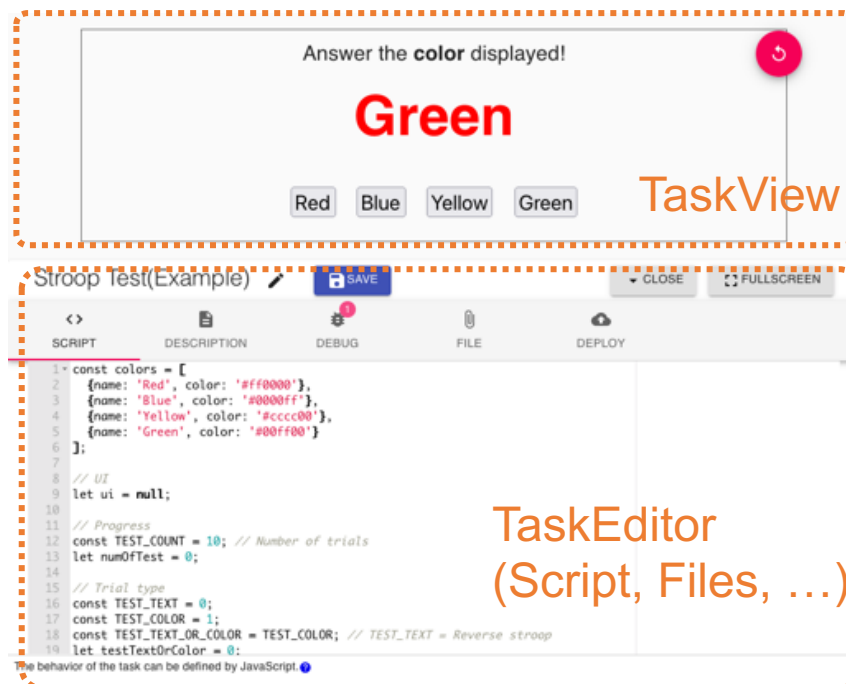


図 3 TaskEditor によるタスクの定義

GO-E-MON を利用する実験者および実験協力者は、Personary との連携管理、実験用タスクの管理のため、自身のアカウントを GO-E-MON に作成する必要がある。GO-E-MON にログインした利用者は誰でも、新しいタスクを作成できる。つまり、利用者は誰でも実験者になることができる。実験者は、Web ブラウザで動作する TaskEditor を用いて、実験タスクの動作を定義するためのスクリプトを編集し、その動作をすぐに TaskView で確認できる。タスクが完成したら、実験者はその実行用 URL を実験協力者に提供することで、タスクの実施を依頼することができる。実験協力者がその URL を開くと、GO-E-MON にログインした状態でなければログインを要求される。実験協力者によるログインがなされると、TaskView は URL 中に指定されたタスク ID に対応するスクリプトやファイルを読み込み、タスクの実行を開始する。

タスクスクリプトは、Web ブラウザ上の TaskView 上のユーザインタフェースの構築処理と、ユーザインタフェースで発生するイベント処理を行うイベントハンドラ、通知メッセージの作成・送信処理やメッセージ受信時の処理のようなサーバースクリプトから構成される。タスクの動作は任意の JavaScript を用いて定義することができ、実験者は Web ブラウザが提供する機能を用いて柔軟にインタフェースを定義できる。TaskView はタスクスクリプトに対して Application Programming Interface (API) を提供する。タスクスクリプトはこの API を用いて、GO-E-MON に対するログの送信指示や、現在の利用者に関する設定情報の取得要求の発行をすることができる。

TaskView が提供する API は、Garmin Connect(Garmin Ltd., 2021)のようなウェアラブルセンサーサービスや、LINE(LINE Corporation, 2011)のようなメッセージングサービスとの連携機能をタスクスクリプトに対して提供する。タスクスクリプトが API を通じて、External Service Connector に対して外部サービスとの連携を指示すると、利用者にはその外部サービスが提供する認可画面が表示される。利用者がこ

の認可画面から、外部サービスと GO-E-MON の連携を許可すると、外部サービスは、サービスにアクセスするための一時的なアクセストークンや、利用者の識別子等の情報を GO-E-MON に対して送信する。GO-E-MON と外部サービスの連携をスムーズに行うためには、これらの情報は素早く取り出せるようにしておく必要がある。そのため、GO-E-MON はこれらの情報を Personary には保存せず、GO-E-MON が持つデータストア User Session DB に保存する。

ウェアラブルセンサーのデータが実験協力者から同期されたり、あるいはメッセージングサービスを介して GO-E-MON に対して実験協力者からメッセージが送信されたりすると、GO-E-MON サーバー上でスクリプトが実行される。実験者は任意のサーバースクリプトを指定することができる。サーバーが持つ計算資源の悪用を防止するため、このスクリプトは 30 秒でタイムアウトするものとし、外部への通信は禁止する。サーバースクリプトは GO-E-MON の API により可能な操作のみが許可される。

タスクから GO-E-MON の API を通じてログ送信が指示されると、そのデータが ResultStore に送信される。ResultStore はそのデータに仮名 ID を付与し、一時バッファにデータを保存した上で、実験者および実験協力者の Personary へと送信する。送信が成功すればバッファからそのデータを削除する。

実験者及び実験協力者は Personary アプリから保存されたデータを確認することができる。(図 4)



図 4 Personary によるログの確認

ResultStore は Personary にデータを書き込むために、GO-E-MON システム専用の Personary アカウント(GO-E-MON システムアカウント)を持つ。利用者は事前に Personary から、GO-E-MON システムアカウントを「友達」として登録する。すると、利用者の Personary 上に、GO-E-MON システムアカウントからのデータを保存するための専用のチャンネルが自動的に生成される。これにより GO-E-MON は各利用者の GO-E-MON 専用チャンネルに実験データを書き込むことができるようになる。なお、利用者が持つ GO-E-MON チャンネル以外のチャンネルは、GO-E-MON システムアカウントからは参照することはできない。

利用者が GO-E-MON システムアカウントと友達になっていない場合、GO-E-MON は ResultStore の一時バッファにデータを保持したまま、次の送信機会を待つ。利用者がチャンネルを共有したことを検出したのち実験データをそのチャンネルに対して保存する。Personary がチャンネルに対して行う操作は新しいデータの追記のみであり、既存データの読み込みや更新は行わない。また、実験者や実験協力者はいつでも GO-E-MON システムアカウントとの友達登録を解除することで、GO-E-MON システムアカウントが自身のチャンネルにデータを書き出すことを停止することができる。

2.3 GO-E-MON の利用者管理

GO-E-MON は Personary アカウントとの関連づけと実験タスクの編集権限管理のため、実験者と実験協力者の双方に GO-E-MON アカウントの作成と、ログインを要求する。実験者は、実験タスクの設定として、実験で得られたデータを受け取るか否かを選択することができる。データを受け取ることが選択された場合、ResultStore は GO-E-MON によって生成された仮名 ID を実験データに付加して、実験者の Personary にデータを送信する。実験協力者があるタスクを初めて開いた際に、GO-E-MON はタスク ID とユーザー ID の組に対してユニークになるようなランダムな ID を生成し、これを仮名 ID とする。この仮名 ID は GO-E-MON のユーザー情報を保持するデータストアに格納され、GO-E-MON の処理時にすぐに参照できるようになる。この仮名 ID により、実験者の Personary チャンネルのデータにはメールアドレスやユーザー名などの GO-E-MON 外で識別可能な識別子を付与することなく、その実験の範囲内で実験協力者を識別することができる。

また、実験を行う上で重要なのは実験群・統制群等の群わけの管理である。GO-E-MON はランダムな群わけを行う機能をサービスとして提供する。実験タスクの設定に、群ラベルをリスト形式で定義しておく、GO-E-MON はその実験タスクを開いた協力者に対して群ラベルをランダムに割り付ける。例えば群ラベルが A, B, C とリスト指定されていた場合は、ある実験協力者には A、別の実験協力者には B あるいは C、といった形でラベルを設定する。実験協力者がそのタスクを初めて開いた際に群ラベルの付与を行う。この処理では、まず群ラベルごとのユーザー数を集計し、最もユーザー数が少ない群ラベルからランダムに群ラベルを選択し、実験協力者に対して割り当てる。決定された群ラベルは仮名 ID と同様に、ユーザー情報を保持するデータストアに保存される。現在のタスクにおける群ラベルは、タスクスクリプトから GO-E-MON API を通じて参照することができる。スクリプトがこのラベルに応じて振る舞いを変えることで、実験的な操作を実現することができる。また、群ラベルが定義された場合、Personary に送信される実験結果データには、仮名 ID に加え、群ラベルが付与される。これにより実験者は群ごとの振る舞いの違いを、仮名化されたデータのみから分析することができる。

2.4 GO-E-MON のタスク定義

実験タスクの振る舞いを定義するタスクスクリプトは, TaskEditor により記述することができる. また TaskEditor はファイルアップローダーを備え, JavaScript を含む任意のファイルをアップロードすることもできる. JavaScript の知識が必要にはなるものの, 実験者は既存の強力なライブラリ, 例えば視線追跡を行う WebGazer (Papoutsaki et al., 2016)や Bluetooth デバイスとの連携(Web Bluetooth Community Group, 2021)ライブラリを活用することができる. このファイルアップローダを利用して jsPsych (de Leeuw, 2015)や WebBCI(Stegman, Crawford, & Gray, 2018)などの JavaScript で書かれたフレームワークをアップロードして利用することも可能である. また, JavaScript ファイル以外にも, スタイルシート (World Wide Web Consortium, 1996)や, 画像ファイルなどのリソースをアップロードすることが可能である. アップロードされたファイルは GO-E-MON のストレージに保存され, 高速に配信可能になる.

タスクスクリプトでは, Personary へのログ保存等の GO-E-MON に特有の処理を, GO-E-MON API を通じて簡単に利用することができる. 例えば jsPsych でタスクを記述する場合, 実験データの記録指示は以下のように記述できる.

```
// 青字コメント: 結果を送信するため, on_finish ハンドラには以下のようなコードを記述する.
// 出力内容は [デバッグ] タブで確認できる.
// `context.finish('人間が読める結果サマリ', 保存するデータ)`
on_finish: function(data) {
    context.finish(`回答を受け付けました`, {
        data: data.values(),
        variables: jsPsych.getAllTimelineVariables(),
    });
}
```

jsPsych では, 定義したタスク(タイムライン)が完了したときに, on_finish 関数というものが呼び出される. このスクリプトでは, on_finish 関数の処理として, GO-E-MON API が提供する context オブジェクトが持つ finish 関数を, Personary に保存したい内容とともに呼び出している. このスクリプトでは, 保存すべきデータとして, jsPsych が保持しているタイムラインにまつわる全てのデータ(getAllTimelineVariables 関数の戻り値)を指定している. この呼び出しにより GO-E-MON はタスクの終了を認識し, finish 関数に指示されたデータを, 実験者および現在タスクを操作している実験協力者の Personary に送る処理を実行する. GO-E-MON を使わず, 独自のコードで Personary にデータを送ろうとすると, 誰の, どのチャンネルに, どんなデータを送らなければならないのかを全て指定しなければならない. GO-E-MON のタスクスクリプトでは, 実験者と実験協力者の各チャンネルは明らかであるので, 誰に送るか, どのチャンネルに送るのか等の手順の指定を行うことなく, 関数の呼び出し一つで Personary に対してデータを送信することができる.

2.4.1 LTI 連携の実現

実験タスクからは、LTI tool provider として実装された教材の呼び出しも可能である。タスク設定に接続先 LTI tool provider の情報(URL, Client ID, Deployment ID 等)を設定し、タスクスクリプトから GO-E-MON API に対して LTI サービスへの接続を指示すると、設定された URL, Client ID, Deployment ID, 認証エンドポイントの情報に従って、利用者を LTI tool provider の画面へと転送し、認証を経た上で実験タスクの画面に LTI tool provider を埋め込むことができる。LTI tool provider との連携は、JavaScript では以下のように記述できる。

```
context.lti.getLoginURL({
  params: params,
}, function(url, error) {
  if (error) {
    console.error('Error: ' + error);
    return;
  }
  const iframe = $('<iframe></iframe>').attr('src', url);
  iframe.css({
    position: 'absolute', width: '100%', height: '100%', left: 0, top: 0,
  });
  context.lti.onGradeUpdate(function(grade) {
    console.log('Grade Changed', grade);
    const score = Object.values(grade.scores)[0];
    const log = JSON.parse(score.comment).log;
    log.data = JSON.parse(log.data);
    ...
    context.finish('LTI tool を実施しました.', {
      scoreMaximum: score.scoreMaximum,
      scoreGiven: score.scoreGiven,
      log: log,
    });
  });
  context.root.append(iframe);
});
```

LTI との連携も GO-E-MON が提供する context オブジェクトを利用する。context.lti.getLoginURL 関数を呼び出すと、タスク設定に記述された連携先の LTI tool provider の接続情報にしたがって、LTI tool provider にログインするための URL が発行される。この URL を表示するブラウザウィンドウを定義すれば、そのウィンドウ内に LTI tool provider の画面を埋め込むことができる。LTI tool provider から送信された成績データは、context.lti.onGradeUpdate 関数で監視することができる。LTI tool provider 側でスコアの送信(Grade の更新)が行われると、GO-E-MON はこれを検知してユーザーが定義したハンドラを実行することができる。この例では LTI tool provider からのスコアの送信を検出したら、それをこのタスクの実験結果データとして Personary に送信している。

LTI tool provider と GO-E-MON の連携フローを以下に示す(図 5)。

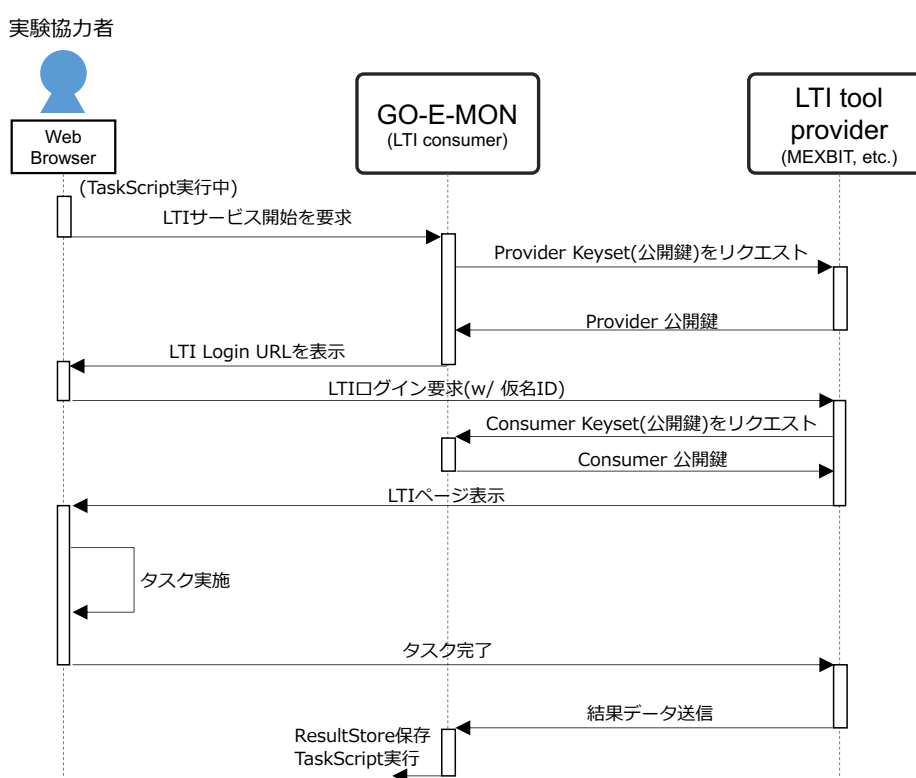


図 5 LTI tool provider との連携フロー

GO-E-MON は LTI consumer として秘密鍵・公開鍵のペアを保持しており、LTI tool provider が公開する公開鍵と自身の秘密鍵を利用しながら、互いに情報を署名・暗号化を行うことで互いに連携する。GO-E-MON の秘密鍵は GO-E-MON から LTI tool provider に引き渡すユーザー識別子等のアクセス情報の署名に用いられる。LTI tool provider に渡すユーザー識別子は実験データに付与する仮名 ID と同様の方法で生成する。仮名 ID に基づいた連携であるため、LTI tool provider からは、GO-E-MON の利用者が具体的に誰なのかを推測することはできない。この情報を受けて、LTI tool provider 側は LTI consumer として GO-E-MON が公開する公開鍵を用いてアクセス情報を検証し、GO-E-MON が提供した識別子であることを確認した上でサービスを提供することで、安全なサービス間の連携が確立される。

また、GO-E-MON は LTI tool provider に対して LTI Grade API(1EdTech Consortium, 2019)を提供する。LTI tool provider において利用者が得た成績に関する情報を Grade API を通じて GO-E-MON へとフィードバックすることができる。Grade API を介して LTI tool provider から送信された情報は、先に示したコードのように、ブラウザで動作しているタスクスクリプトで処理することができる。このような複雑な連携フローを GO-E-MON がサービスとして提供することで、実験者はこれらの知識がなくとも簡単に、LTI tool provider に準拠した学習教材との連携タスクを実現することができる。

2.4.2 OAuth 連携の実現

GO-E-MON API は LTI Tool Provider 以外にも OAuth 2.0(Internet Engineering Task Force, 2012)に準拠したシステムとの連携機能を提供する。GO-E-MON では OAuth 2.0 で連携可能な外部サービスとして、以下の 2 種類のサービスをサポートする。

- メッセージングサービス: 一例として LINE に対応
- ウェアラブルセンサーサービス: 一例として Garmin Connect に対応

外部サービスとの連携は、GO-E-MON JavaScript では以下のように記述できる。以下は LINE のメッセージで実施のリマインドを行うためのコード例である。

```
context.messaging.request({
  type: 'line'
}, function(result, error) {
  var script = "if (source) { メッセージ受信時の処理(source); } else { 定期処理(); } ";
  // 実験協力者のタイムゾーン 7 時-21 時の間, 毎時間 (毎時 0 分に) 実施する
  context.messaging.setSchedule({
    schedule: '0 7-21 * * *',
    script: script,
  });
});
```

メッセージングサービスおよびウェアラブルセンサーサービスとの連携開始は、LTI と同様に、GO-E-MON が提供する context オブジェクトから要求する。context.messaging.request メソッドの呼び出しにより、利用者に対して LINE へのログインと、LINE アカウントと GO-E-MON 連携の可否を問い合わせる画面が表示される。利用者はこれを拒否することができ、実験タスクは拒否された場合に対する応答を JavaScript 内で適切に記述しなければならない。タスクスクリプトは利用者から LINE サービス利用の認可を得た場合の処理として、context.messaging.setSchedule 関数により定期実行間隔と、実行すべきサーバースクリプトを指定することができる。スクリプトはサーバ側で指定された時刻またはメッセージ受信時に実行されるため、実験実施者は必要な処理をスクリプト中に記述することで、定期的なリマイン

ドやメッセージに対する応答処理を記述することができる。サーバースクリプトは、悪意のあるコードを GO-E-MON サーバー上で実行できないよう、JavaScript Sandbox 内で実行する。

外部サービスは OAuth 2.0 に準拠した認可機構を持つものを前提とする。連携フローの例を以下に示す(図 6)。

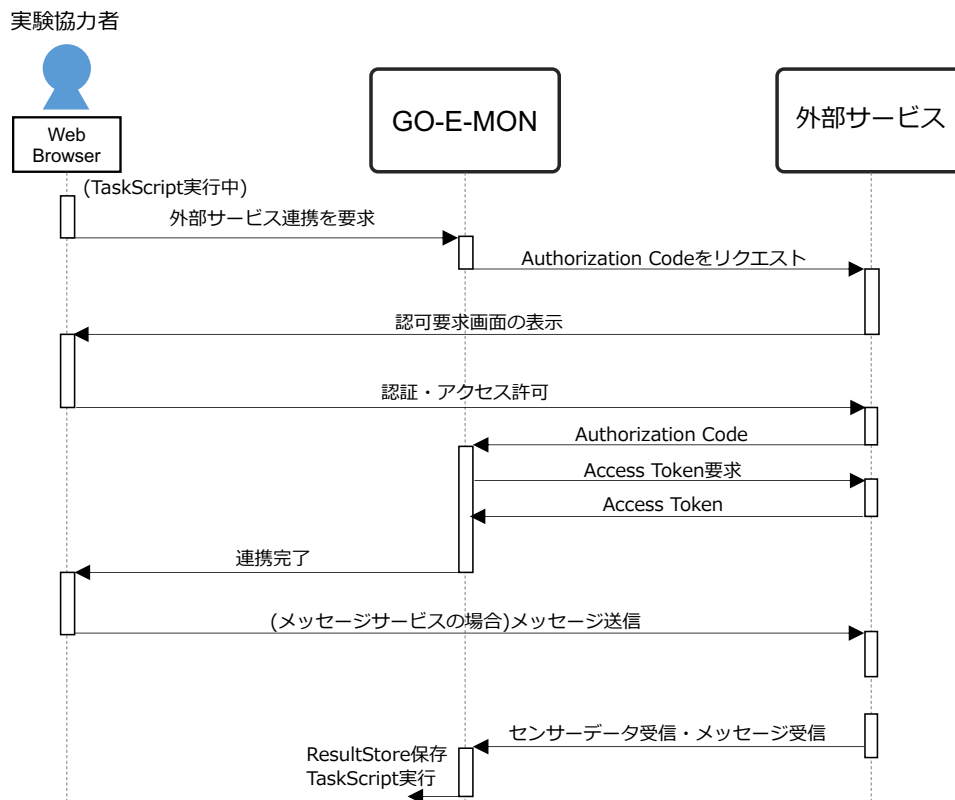


図 6 外部システムとの連携フロー

GO-E-MON は連携対象のサービスに対して OAuth Client として振る舞う。GO-E-MON が各サービスからあらかじめ得た OAuth Client 情報に基づき、外部システムと OAuth Session を確立し、実験協力者に対してサービスアクセスの認可を要求する。実験協力者が認可を選択すると、外部システムにアクセスするためのアクセストークンや、実験協力者の識別情報が、外部システムから GO-E-MON に対して提供される。GO-E-MON はこれらの情報を User Session DB に保存し、その後のタスク実行に利用する。

アクセストークンは機微情報であるが、一般的に OAuth 2.0 サービスのアクセストークンは任意のタイミングで無効化することができるため、漏洩しても損害を最小限で済ませることができる。また、実験協力者の識別情報は OAuth Client ごとに異なる識別子が使われているため、その識別子だけでは名寄せなどに利用できないよう、外部サービス側で配慮されている。このように OAuth 2.0 で得られる情報には安全性に対して配慮がなされているが、これらの情報が漏洩することは好ましくない。そのため、User Session DB の保持期間を 2 週間とし、それ以降は削除することで漏洩の危険性を抑える。

長期にわたって User Session DB の内容を維持したい場合は, 実験協力者を実験タスクへと定期的に誘導し, 実験協力者の同意のもとこれらのデータを維持できるものとする. 実験協力者が実験参加をやめれば, 一定期間経過後に User Session DB のデータは削除される.

3. GO-E-MON の実装

GO-E-MON の設計に基づき, GO-E-MON の実装をおこなった. GO-E-MON は URL <https://goemon.cloud> で試験公開している. 実装した機能と利用サービス・技術の関係を示す(表 2).

表 2 GO-E-MON の構成技術

機能	利用した技術
TaskView, TaskEditor	Firebase Hosting / React.js
ResultStore	Cloud Functions for Firebase / Cloud Firestore / Personary
Task DB	Cloud Firestore / Cloud Storage for Firebase
User Session DB	Cloud Firestore / Cloud Functions for Firebase
External Service Connector (LINE)	Cloud Functions for Firebase
External Service Connector (Garmin Connect)	Cloud Functions for Firebase
External Service Connector (LTI)	Cloud Functions for Firebase

全ての機能は Google 社が提供する Firebase(Google LLC, 2012)に構築した. 実験タスクの UI となる TaskView および TaskEditor は, Web アプリケーションのフレームワーク React.js(Meta Platforms, Inc., 2013)を用いて実装し, 作成した HTML ファイル等のリソースファイルを Firebase Hosting に配置した. Firebase Hosting は高速化のためにファイルを Content Delivery Network (CDN) (Zolfaghari et al., 2020)にキャッシュする. CDN にファイルをキャッシュすることで, 多数のアクセスに対して高速に応答することができる. また, ResultStore は Cloud Functions for Firebase として 5 分間ごとのバッチ処理として実装した. 一時保存領域として Cloud Firestore を用いることで, 一時的なデータの高速な格納を実現しながら, 安全な PDS に対してログデータを書き込むことで, 安全性と性能の両立を狙った. また, タスクスクリプト等のタスク設定情報は高速にアクセスできるよう Cloud Firestore に, 実験者によりアップロードされたファイルは Cloud Storage for Firebase に保存する. Cloud Storage からの取り出し時は TaskView, TaskEditor と同様に CDN を有効化することで, ファイルをキャッシュし読み込みを高速化する. Task DB に保存されたファイルは, ブラウザのキャッシュに保存され再利用されるか, CDN にキャッシュされることで高速にアクセスができる. User Session DB は Cloud Firestore で実装する. Cloud Functions for Firebase により期限切れ処理を定期実行することで, 古いトークン等の情報を削除する. また, External Service Connector としてメッセージング API として LINE, 生理指標収集 API として Garmin Connect, アプリケーション連携機能として LTI 連携機能を実装した.

4. まとめ

本章では Real World を指向したインタラクション研究を実施可能な実験環境として、オンライン実験プラットフォーム GO-E-MON の設計、実装を行った。

各種デバイスにおける Web ブラウザの高度化だけでなく、様々なサービスが Web 関連技術により接続可能になることで、Web アプリケーションによるオンライン実験の柔軟性は非常に高くなった。しかしサーバーの構築や運用、外部サービスとの接続は多くの実験者にとっては負担となる作業であり、その負担を軽減する必要がある。また、長期にわたって途切れなくサービスを提供しようとする、高い可用性を維持するためのノウハウも必要となる。そこで GO-E-MON がサーバーの管理を担い、各種機能をサービスとして提供することで、可用性と簡便性を実現する。

Web アプリケーションはインターネットに接続されたサーバーによって提供するため、必ず集中的に管理せねばならない部分が残る。これらのサーバー群の管理には、侵入等の攻撃とそれによるデータ漏洩のリスクへの配慮が常に必要となる。そこで GO-E-MON は、データ管理に分散型 PDS である Personary を組み合わせることで、実験データが GO-E-MON サーバー内に保持される期間を最小化し、長期的には実験データを各実験者のストレージにて分散管理させることで安全性を確保する。ただし、アプリケーション連携のための情報などの、長期に維持する必要はなく、かつ高速な読み書きを必要とするデータについては、クラウドが提供する高速なデータストアに保存し、適切に期限切れ処理を行うことで安全性と高速な動作を両立する。

また、実験協力者の Personary に実験結果データを送信することで、実験協力者が自身の意志に基づき、自身の実験結果データを利用することができる。実験協力者によるデータ利用が可能になることで、従来の実験者のみがデータを管理する方法では困難であった、複数の実験者による実験データの統合が可能になる可能性がある。

第 3 章 実験プラットフォーム GO-E-MON の評価

1. 評価実験の概要

GO-E-MON が Real World 認知科学を指向したオンライン実験環境の特性を持つことの評価のため、4 つの実験を事例として示す。各実験と評価対象を示す(表 3)。

表 3 評価実験と評価対象

No.	実験名	対象・ 人数	期間	実験タスク	柔軟 性	安全 性	簡便 性	可 用 性	流 通 性
1	PBL 型講義での試 用	大学生 19 名	3 ヶ月	講義動画プレイヤー 数字記憶タスク		✓	✓		✓
2	講義動画による自学 自習	高校生 78 名	1-2 時間	講義動画プレイヤー	✓	✓	✓		
3-1	学習アプリケーション による自学自習	高校生 1422 名	3 ヶ月	マンガリッシュ 英単語クイズ 自由英作文	✓	✓	✓	✓	
3-2	学習アプリケーション による自学自習 + 生理指標収集	大学生 20 名	1 ヶ月	マンガリッシュ 自由英作文	✓	✓	✓		

最初に、実験者と実験協力者が GO-E-MON と Personary の操作を理解し、使用することができるかを検証することを目的とし、大学生による PBL 型講義での試用を行った(実験 1)。実験 1 では、特に GO-E-MON で得られた実験データを Personary に格納することで、安全性と流通性が得られるか、実験者がシステム運用の手間を負うことなく実験ができる簡便性が得られることを確認する。

以降は、GO-E-MON が可能なことや課題を示すための事例として、実際に GO-E-MON を用いて行われた学習研究に関する実験について述べる。GO-E-MON が柔軟にタスクを構成できることを示す事例の一つとして、実験 2 を紹介する。実験 2 は、自学自習の場面における、講義動画の視聴行動を、YouTub 動画プレイヤーを活用した講義動画プレイヤーを用いて収集する実験である。実験 2 では、実験者が GO-E-MON 開発者から提供されたサンプルを用いて、柔軟にタスクを実装できること、実験をシステム運用の手間を負うことなく実施できること、GO-E-MON はデータを集中管理せず、実験者の Personary により管理できる安全性を備えることを示す。

GO-E-MON がウェアラブルセンサー等の外部システムと連携する柔軟性を持ち、長期的な(数日～数ヶ月の)学習教材に対する多くの学習者の行動データを、同時に、継続的に計測できることを示す事例として、実験 3-1 と実験 3-2 を紹介する。実験 3-1 は高校生の、実験 3-2 は大学生の協力を得て行

われた実験である。実験 3-1 では、実験 2 と同様の柔軟性と簡便性、安全性に加え、1000 名を超える多くのユーザーの利用に、数ヶ月にわたって GO-E-MON が耐えることができる可用性があることを確認する。実験 3-2 では柔軟性を示す事例として、ウェアラブルセンサーによる生理指標データの同時収集も行い、柔軟に外部システムとの連携が実現できることを確認する。

2. 実験 1: 大学生による PBL 型授業での使用

2.1 実験の概要

実験者と実験協力者が GO-E-MON と Personary の操作を理解し、使用することができるかどうかを検証するため、東京大学の学部 1 年生を対象とした「初年次セミナー」を選択した学生 19 名に協力してもらい、GO-E-MON を用いた実験を実施する。

本実験では、2 種類の状況を設定して実験を行なう。

- (1) 1 人の実験者が複数の実験協力者に対して実験タスクを提供する状況
- (2) 実験協力者が得た自身のデータを実験者以外に提供する状況

まず(1)では、東京大学開研究室の修士課程学生 A が実験者となり、GO-E-MON を使い実験協力者の動画視聴行動を収集する実験を行う。COVID-19 の流行時にオンライン講義の割合が増え、映像による講義を主体とした MOOCs (Massive Open Online Courses) の利用が急増していることが報告されている(Impey & Formanek, 2021)が、登録者数の多い MOOCs の修了率は 10%程度という統計もある(Jordan, 2015)。動画教材において学習継続を阻害している要因がないかどうか検証することが重要である(Kim et al., 2014; Liang, Li, & Zheng, 2016)。講義動画に対する学習者の視聴行動の詳細な記録は、この課題解決に必要とされる可能性がある。そこで本実験では、GO-E-MON を使って動画視聴行動記録タスクを用いた講義を実施することを通して、Personary を用いて実験者としてデータの集約が可能かどうかの確認を行う。

次に(2)として、グループワーク型の演習として、参加者が自分自身あるいはグループメンバーの Personary に格納された実験データを分析する実験を行う。本実験では参加者は「記憶と脳活動」をテーマにグループ実験をすることが指示される。参加者は実験用タスクとして、(1)の動画視聴行動収集タスクに加え、数字記憶課題を実装したタスクが提供される。また、Garmin vivosmart 4 ウェアラブルセンサー(Garmin Japan, 2021)、MUSE-Classic 携帯型脳波計(InteraXon, 2021)が貸与され、これらのデータの解析方法が案内される。学生自身が実験協力者としてタスクを実施し、自身のデータをグループメンバーに開示、解析し、プレゼンテーションを行う形式とすることで、実験協力者が得たデータを本人同意のもと利用する際の課題について調査する。実施には、データを学生が分析してプレゼンテーションを行うグループ講義の例(Donoghue, Voytek, & Ellis, 2021; Millman et al., 2018)等を参考にした。

2.2 実験方法

本実験は東京大学の学部1年生を対象とした「初年次セミナー」内で実施する。本講義は「人間の記憶」をテーマとした講義であり、その一環として本実験を実施する形とする。講義は以下の2つのパートにより構成される。

- (1) 動画視聴型講義における GO-E-MON によるデータ収集実験
- (2) グループワーク型講義における GO-E-MON を利用したデータ取得・共有実験

(1)は記憶術をテーマとした3つの講義の動画を別々の実験タスクとして提供し、各タスクの URL を学生に配布するものとする。動画はメモリスポーツと呼ばれる、記憶をテーマとした競技において活動をしている実験実施者 A が作成する。動画視聴講義の後、(2)として4~5人の4つのグループに分かれてグループワークを行う。実験協力者には、実験用タスクと、ウェアラブルセンサーGarmin vivosmart 4、脳波計 MUSE-Classic が実験者から提供される。実験データの取得方法と解析方法を説明する講義を2回行ったのち、1週間後に最終発表会を行う。

2.2.1 実験用タスク

本実験では参加者が利用可能な実験タスクとして、2つの GO-E-MON タスクを用意する。1つは動画プレイヤーを表示し、その視聴行動を記録するタスク、もう1つは数字を記憶するタスクである。動画配信サーバーを用意することは手間がかかるため、講義動画配信にはYouTubeを活用する。動画はYouTubeに限定公開としてアップロードするものとし、組み込み型YouTubeプレイヤーをGO-E-MON タスク内に組み込む形で実装する。このプレイヤー上での視聴行動はYouTube Iframe API (Google Developers, 2018) を用いて記録する。動画プレイヤーの再生状態の変化はGO-E-MON API を通じて実験結果データとしてPersonaryに記録される。記録されたデータはGO-E-MON のResultStoreに保存され、実験者および実験協力者のPersonaryへと送信される。

動画視聴タスクのユーザーインターフェースおよびログ形式を示す(図7)。

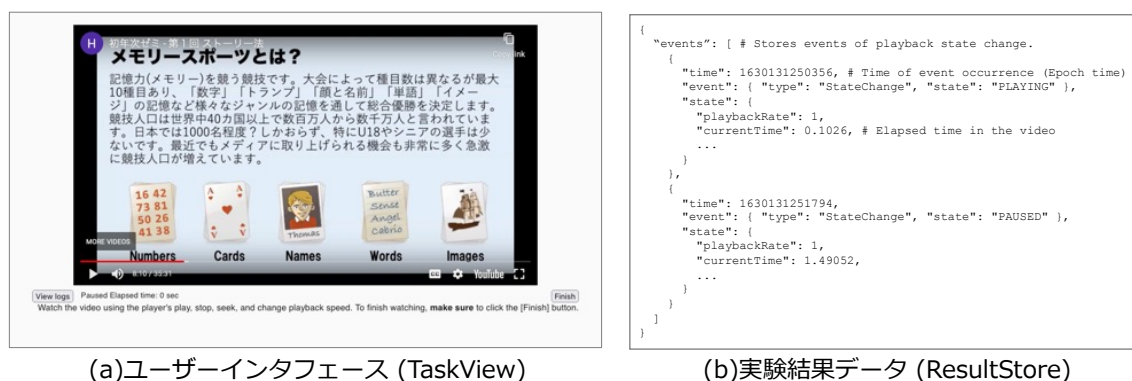


図7 動画視聴タスクのユーザーインターフェースとログ形式

動画視聴タスクのユーザーインターフェース(図7(a))は、YouTubeの動画プレイヤーに準拠する。学習者が馴染み深いインターフェースにより、再生速度変更やスキップなどの操作を行いやすくする。

動画視聴タスクは以下のような JavaScript コードにより定義される.

```
Const videoId = 'XXXXXXXX';

player = new YT.Player('player', {
  height: '360', width: '640', videoId,
  events: {
    'onStateChange': onPlayerStateChange,
    ...
  }
});

// ユーザーインターフェースの作成

const finishButton = $('<button>視聴終了</button>');

finishButton.click(() => {
  ...

  context.finish(`動画を視聴しました`, {
    history,
  });
});

...

// プレイヤー状態変化の検出

function onPlayerStateChange(event) {
  ...

  history.push({
    time: Date.now(),
    event: {
      type: 'StateChange',
      state: toPlayerState(event.data),
    },
  });
  ...
}
```

動画プレイヤーは YouTube Iframe API が提供する YT.Player コンストラクタを利用して作成する. YT.Player には, onStateChange ハンドラが用意されており, 利用者による操作が行われるとプレイヤー状態が変化し, この関数が呼び出される. この呼び出し記録をタイムスタンプとともに変数に格納しておき, 「視聴終了」ボタンが押されたら GO-E-MON の context オブジェクトを介して操作記録を送信し, タ

スクとしては終了する。プレイヤーの状態変化の記録は図 7(b)のように、どのような変化が、いつ起こり、その時のプレイヤーはどのような状態だったのかといった情報が構成される。このスクリプトを GO-E-MON 開発者がサンプルとして作成し、実験者に提供した。実験者は提供されたサンプルに基づき、動画の ID を自身がアップロードした動画のものに差し替えるなどの変更を行なった上で動画視聴実験を実施した。

数字記憶タスクは 60 桁の数字を記憶し、キーボードやタッチ操作で回答する形式のタスクである。タスクは動画タスク同様に JavaScript で実装し、ユーザが入力するデータを制御するために、jQuery (Open JS Foundation, 2006) ライブラリを使用した。

数字記憶タスクのユーザーインターフェースおよびログの形式を示す(図 8)。

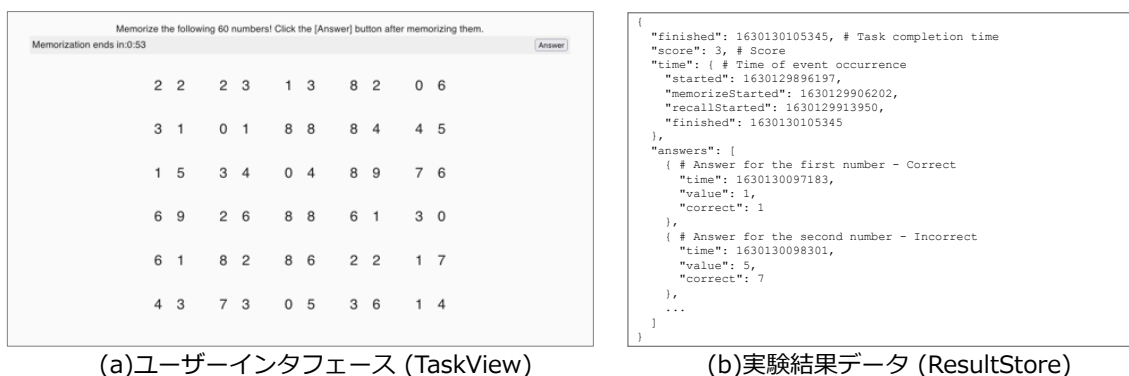


図 8 数字記憶タスクのユーザーインターフェースとログ形式

数字記憶タスクは 60 桁の数字を特定の時間内に繰り返し記憶するものである。記憶フェーズでは記憶すべき数字列が表示され(図 8(a))、回答フェーズでその数字列を学習者がキーボード等で回答するという形式をとる。記憶させる数列はランダムに生成され、解答内容、正解・不正解、解答のタイミングがタイムスタンプとともに記録される(図 8(b))。

実験協力者には、タスクを実行する前に、大学の Google Workspace アカウントを使用して Personary アカウントを設定することを依頼する。その後、GO-E-MON システム Personary アカウントを友達として追加するよう依頼する。これにより GO-E-MON タスクで得られた実験データが実験協力者の Personary チャンネルにも保存されるようになり、後述するデータ解析環境で自分自身のデータを解析することができるようにする。

2 つの実験タスクに加え, ウェアラブルセンサーGarmin vivosmart 4 と脳波計 MUSE-Classic から, 実験協力者がデータを取得するための方法を用意する. 機器とデータフローの関係を示す(図 9).

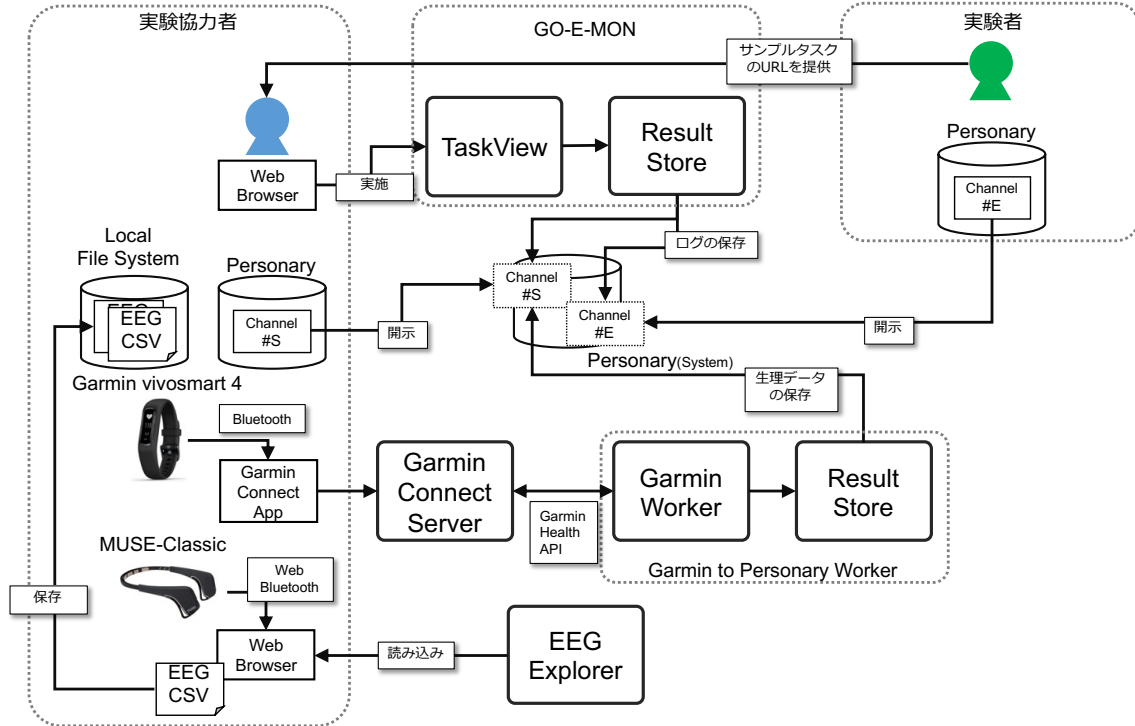


図 9 サンプルタスクのデータフロー

GO-E-MON は Garmin vivosmart 4 から, Garmin Health API(Garmin Ltd., 2021) を通じて, 睡眠記録や脈波センサー値を取得する. 取得されたデータは実験協力者の Personary に保存される. Garmin のデータは実験協力者が自身で分析するために収集するものとし, 実験者の Personary には保存しない. また, MUSE-Classic 脳波計のデータ取得には Web ブラウザで動作する EEG Explorer(Shaked, 2021)を使用する. EEG Explorer は Web Bluetooth API を通じて MUSE-Classic と接続することができる. 取得した脳波データは実験協力者の PC のメモリに保存できる. 実験終了後, このデータを実験協力者のローカルファイルシステムに保存することができる. GO-E-MON のタスクも Web Bluetooth API を介して MUSE-Classic に接続することが技術的には可能であるものの, タスクの開発に十分な期間が取れなかったため, 既存の EEG Explorer を使用するものとした. したがって, GO-E-MON で作成した各タスクの実験データと Garmin vivosmart 4 による各種計測データは実験協力者の Personary チャネルに, MUSE-Classic で生成した脳波データはローカルファイルシステムに格納されるものとする.

2.2.2 データ解析環境

実験参加者にはデータ解析環境として、Web ブラウザで利用可能な JupyterHub(Project jupyter team, 2019)を用いた計算環境を提供する。本環境は実験参加者の大学アカウントでログインできるように設定する。JupyterHub には、Personary アカウントと接続可能な Jupyter Notebook の実行環境をあらかじめ設定しておく。実験協力者は、説明用に用意された Jupyter Notebook 形式(Kluyver et al., 2016)の説明ファイルに記述された指示に従い、自身の Personary アカウントとの接続設定を行うよう指示される(図 10)。

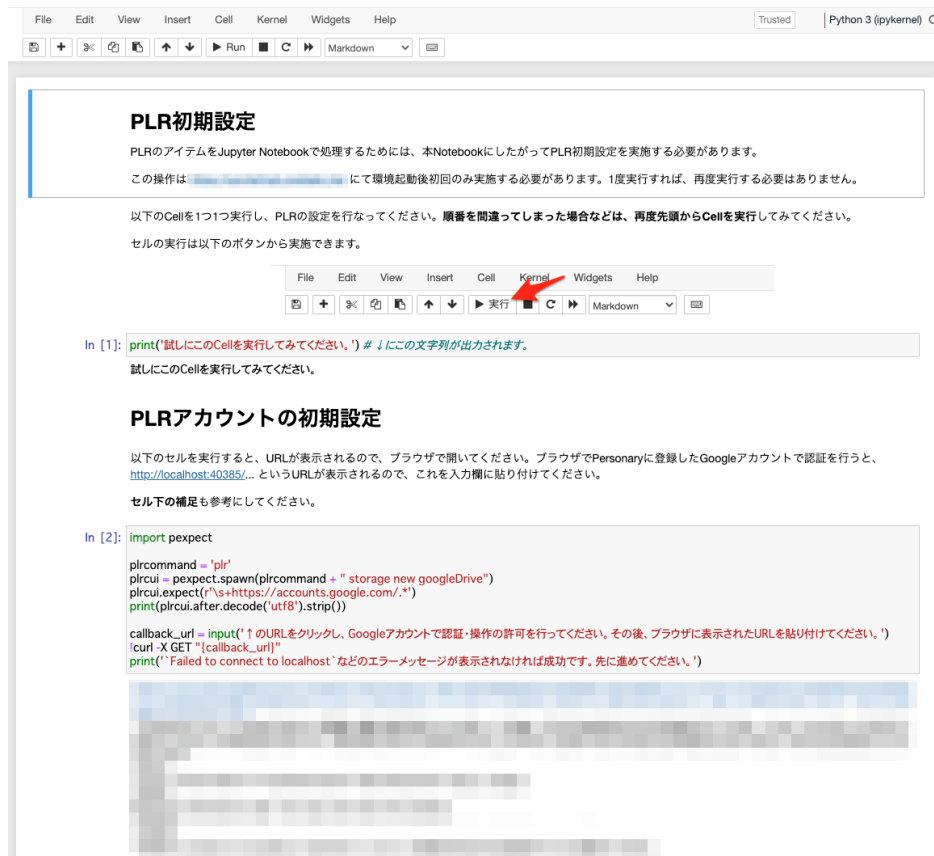


図 10 解析用 Jupyter Notebook の画面例

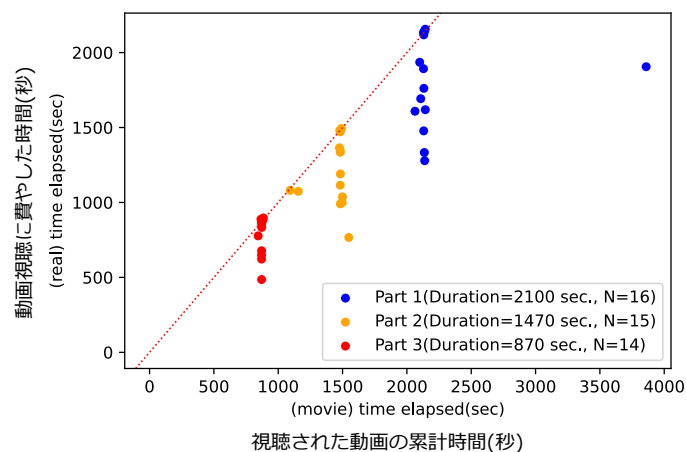
実験協力者は、実験者が用意した Garmin データの分析例を記載した Jupyter Notebook や動画プレイヤーログ、記憶タスクログの分析例を記載した Jupyter Notebook を通して、解析方法を学び、解析を実施する。なお、MUSE-Classic のデータに対する分析手順は、JupyterHub 環境にデータファイルをアップロードして解析することを想定して作成する。解析用の JupyterHub 環境には Numpy(Harris et al., 2020), SciPy(Virtanen et al., 2020), Pandas(McKinney, 2011)をあらかじめインストールして提供することで、各グループがインターネット上に公開されている多くの技術情報を利用しながら、独力で自身のデータを分析できるようにする。

2.3 実験結果

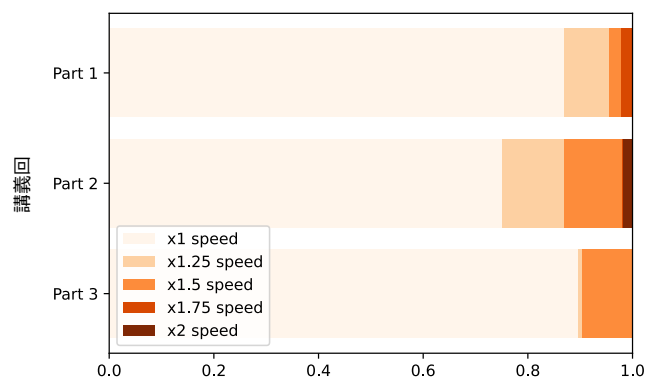
実験 1 は, 東京大学の学部 1 年生を対象とした「初年次セミナー」に参加した大学生 19 名に実験内容の説明を行い, 倫理申請 246-17 に基づき, 学生本人(全員 18 歳以上の大学生)から書面による実験参加への同意を得たうえで実施した。

2.3.1 実験実施者による実験データの収集

まず, (1)動画視聴型講義における GO-E-MON によるデータ収集実験として, 記憶術をテーマとした 3 つの講義の動画プレイヤーを別々の GO-E-MON タスクとして実験協力者に配布し, 各動画を視聴してもらった。その後, GO-E-MON により, 仮名 ID が付与された動画視聴ログが実験者の Personary チャンネルに集約されたことを確認できた。実験者は自身の Personary 中のデータを用いて動画視聴ログを分析することができる。記録された動画視聴ログの概要を示す(図 11)。



(a) 視聴された動画の累計時間とそれにかけた時間



動作視聴中において再生速度が選択された割合

(b) 動画視聴中の再生速度の割合

図 11 動画視聴行動の収集例

グラフ図 11(a)は各実験協力者の動画視聴時間と実際の視聴時間の関係をプロットしたものである。赤線は、視聴された動画の累計時間 = 動画視聴に費やした時間となる直線である。各講義回とも欠席者があり、ビデオを再生した実験協力者の人数は異なる。また、講義回ごとに映像の長さが異なっており、X 軸の値は各回の映像の長さに集中している。しかし、Y 軸の値にはばらつきがあり、シーク機能で再生位置を進めたり、再生速度を変えたりすることで、実際の学習時間が短縮されていることが視聴ログから確認できた。なお、再生速度を変更した場合でも、音声は適切なピッチで再生されるよう YouTube 動画プレイヤーにより制御されている(Powale, 2021)。第 1 回(Part 1)では、ある実験協力者が視聴した映像の総時間が、映像の長さの約 2 倍になっているが、これはこの実験協力者が映像全体を一度早送りして、再度最初から再生したためである。

グラフ図 11(b)は動画視聴時間において、実験協力者がどの再生速度を選択したか、その割合を示したものである。この図から、第 1 回(Part 1)では 1.5 倍以上の再生速度で動画を再生した実験協力者は少なかったが、第 2 回(Part 2)、第 3 回(Part 3)では 1.5 倍以上の再生速度で動画を視聴した実験協力者の割合が増加していることが分かる。スタンフォード大学の Lagunita Platform の視聴行動データを用いた研究(Lang et al., 2020)では、実験協力者 21,835 人のうち動画を再生する際に再生速度を変えたのは 4,345 人であり、再生速度の中央値は 1.25 倍から 1.5 倍であったことが報告されている。本実験結果でも、同じ視聴行動をとった実験協力者が見られた。本実験では、視聴記録から動画全体の再生所要時間のみを可視化することとどめたが、実験 2 では講義動画の提示方法を変化させる介入を行い、学習者の視聴行動がどのように変化するかをより詳細に分析する。

2.3.2 実験協力者自身によるデータの収集・利用ができること

動画視聴講義の後、4～5 人の 4 つのグループに分かれてグループワークを行った。実験データの取得方法と解析方法を説明する講義を 2 回行ったのち、1 週間後に最終発表会を行った。各グループの発表内容と使用したデータの種類の示す(表 4)。

表 4 学生グループによる発表と使用したデータの種類

	動画視聴 タスク	数字記憶 タスク	Garmin Vivosmart 4	MUSE- Classic (脳波) ¹	分析テーマ
グループ 1		✓		✓	起床後および就寝前の脳波と 数字記憶課題の成績との関係
グループ 2		✓	✓	✓	脳波、心拍数と数字記憶課題 における成績との関係
グループ 3				✓	脳波と独自の多桁計算タスク の実行との関係
グループ 4			✓		睡眠とストレスの関係

¹ 脳波データのみ PC にローカルに保存し、その他のデータは GO-E-MON を通じて Personary に保存し、グループメンバーで共有した。

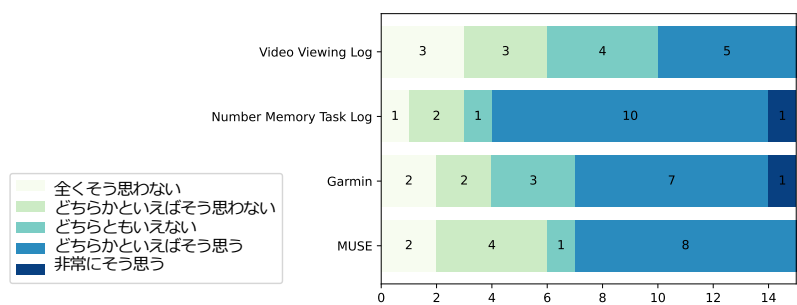
3 班を除く各班は、Personary に保存されたデータを解析する課題を実施しており、Personary に格納された Garmin データおよび数字記憶タスクの実験データが参加者により解析できていることを確認することができた。今回はグループワークとして実施したため、実験協力者全員がデータを解析できたとは限らず、プログラミングやデータ解析が得意な学生の貢献が大きいと思われるが、Personary を用いて、実験協力者が自らのデータを解析することが技術的に可能であることを確認できた。

最後の講義の終了時に、各参加者に対してアンケートを実施した(表 5)。

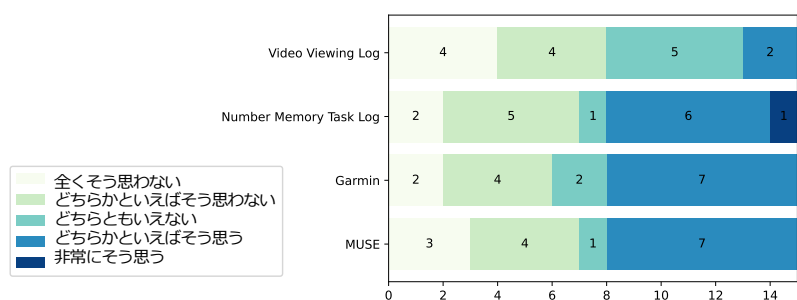
表 5 講義後のアンケート

No.	質問	回答
1	今回の講義で、自身のデータ分析ができたと感じますか。	0: 全くそう思わない 1: どちらかといえばそう思わない
2	今回の講義で、グループメンバーのデータを分析できたと感じますか。	2: どちらともいえない 3: どちらかといえばそう思う
3	この講義で得られた各データについて、理解できたと感じますか。	4: 非常にそう思う
4	この講義で得られたデータを重要だと思いますか。	

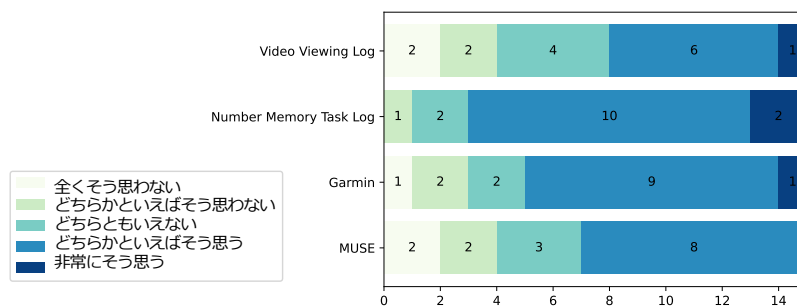
各質問はデータの種類ごと(動画視聴ログ(Video Viewing Log), 数字記憶タスクログ(Number Memory Task Log), Garmin データ(Garmin), MUSE 脳波データ(MUSE))に対して問う形とした。どちらとも言えないを許容するよう5段階で回答するものとし, 15名の学生から回答を得ることができた。アンケート結果を示す(図 12)。



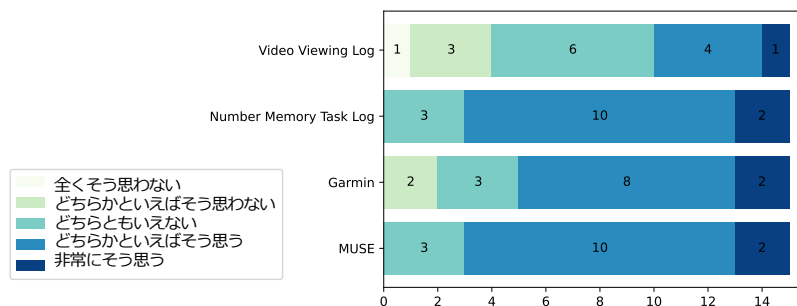
(a)Q1: 今回の講義で、自身のデータ分析ができたと感じますか



(b)Q2: 今回の講義で、グループメンバーのデータを分析できたと感じますか



(c)Q3: この講義で得られた各データについて、理解できたと感じますか



(d)Q4: この講義で得られたデータを重要だと思いますか

図 12 講義後アンケートの集計結果

アンケートの結果、分析できたと感じるか(図 12(a)(b))について、動画視聴ログ以外は「そう思う」以上の回答が多い傾向を持つことが確認できた。これは各グループが動画視聴記録の分析をグループワークのテーマとして選択しなかった(表 4)ことに関係していると考えられる。また、理解できたと感じるか(図 12(c))についてもビデオ視聴記録は数字記憶タスクに比べて「そう思う」以上の回答が少ない傾向が確認できた。

興味深いのは、重要だと思いますか(図 12(d))に対する回答である。動画視聴記録については「そう思う」以上の回答が、他の項目よりも少ない。Personary データの分析方法の説明およびサンプルは、動画視聴記録の分析と、数字記憶タスクの分析の双方に対して提供したが、動画視聴ログに比べて数字記憶タスクのログは正解・不正解と解釈がしやすいデータであったために、数字記憶タスクの方が重要であると感じた可能性がある。分析テーマを参加者が選択するのではなく、動画視聴記録の分析が必要なテーマを参加者に対して実験者から割り当てることで、動画視聴ログの重要性を参加者が認識できる可能性がある。

なお、アンケートに回答した 15 名は全員 Personary チャンネルを共有しても良いと回答したため、その実験協力者と実験者の間で Personary チャンネルの共有作業を実施し、それぞれの実験協力者ごとの実験データが保存されていることを確認した。

2.4 考察

2.4.1 実験実施者チャンネルと実験協力者チャンネルの使い分け

GO-E-MON では実験データの集約方法として、実験者の Personary に対する仮名化済みデータという形での集約と、実験協力者の Personary に保存された実験データの開示による集約の、2 つの方式がある。本実験では、学生が実験者となり、講義の参加者には実験協力者となってもらうことで、これら 2 種類の集約方式が機能していることを確認できた。また、参加者 19 名のうち 15 名は自身の Personary チャンネルの開示を許諾したが、4 名は行わなかった。この 4 名は動画視聴実験でのデータの収集には同意しており、部分的にデータ開示に同意した形となった。このように特定の実験に関するデータ収集だけは同意し、自分自身のチャンネルの開示には応じない参加者は他の実験においても十分考えられるため、実験者によるデータ集約方法は必要なものであると改めて確認できた。

また、今回の実験ではどの参加者も Personary の動作に必要なスマートフォンあるいは PC を持ち、機器の操作にも十分なリテラシーを備えていたため、Personary アカウントを 19 名全員が作成、操作することができた。しかしながら、機器の操作が苦手な参加者や、Google アカウントや Personary アカウントの作成を行いたくない参加者、あるいは乳幼児のようなアカウントが作成可能な年齢に達していないような参加者も想定すべきである。実験協力者が自身でデータを管理することに関しては考察の章で改めて検討する。

2.4.2 実験データの長期管理

講義で使用した動画はそれぞれ異なる実験用タスクとして作成されていたため、GO-E-MON は実験協力者に対して動画ごとに異なる仮名 ID を生成し、実験データに付与をする形で仮名化を行う。本実験では実際に実験者が各動画のログを ID に基づいて突合できないことを確認することができた。また、各実験協力者の動画再生記録は各自の Personary アカウントにも保存されるため、実験協力者の同意によりデータを実験者へと開示することで、各動画のログが突合可能な状態で実験者に提供することができる。データ開示に同意した 15 名については、個人内での講義回ごとの行動の変化を追跡できることもあわせて確認できた。

実験者の Personary を用いてデータ管理を行うということは、その実験者がデータ管理上の全ての責任を負うということも意味する。これはその実験を行なっている最中だけでなく、その後においても同様である。実験者が学生の場合には、大学の Google アカウントを使うのであれば、そのアカウントが在籍期間を過ぎると削除されることを考慮しなければならない。そのため、本研究で実施した実験では、実験終了後、実験実施者の Personary から実験データの取り出しを実施し、別のストレージにて長期にアーカイブする方法を採ったが、Personary でも長期にデータを維持するためのアカウント移行作業などを支援する必要があると考えられる。

2.5 結論

実験 1 では実験者と実験協力者が GO-E-MON と Personary の操作を理解し、使用することができるかを確認するためのケーススタディとして、大学生の PBL 講義を例に GO-E-MON を使った実験を行った。この実験により、GO-E-MON と Personary が提供する各機能を実験参加者が利用でき、分散管理による安全性と、PDS による流通性を得られることが確認できた。実験者は自身の Personary に実験データを仮名化された状態で収集でき、また、実験協力者も自身の Personary に実験データを保存し、チームメンバーに開示し共有することが可能であることが確認できた。

実験者の役割はシステム管理には詳しくない学生が担い、システム管理の負担などは不要であり簡便性を持つことが確認できたものの、実験タスクの作成に関しては GO-E-MON 開発者がサンプルタスクを提供し、この設定部分を書き換えて利用する形となった。そのため、実験タスクを誰でも用意できるわけではないという点で、簡便性に課題が残った。また、実験終了後にアンケートを実施し、実験協力者は理解しやすい形式のデータほど、重要であると感じる可能性が示唆された。これは、分散型 PDS を用いたデータ流通の課題として検討する。

3. 実験 2: 講義動画に対する学習動態の分析

3.1 実験の概要

実験 1 では GO-E-MON と Personary により安全性、流通性が得られるかのケーススタディとして、大学生の講義において GO-E-MON を使用できるかを確認した。次に、GO-E-MON により実現可能なことと課題を示すための事例として、実際に GO-E-MON を用いて行った学習研究を紹介する。実験 2 は講義動画による学習効果の研究のために実施されたものである。一般に講義動画は撮ったものをそのままではなく、適切に編集等を行うべきとされているが、これをする場合としない場合とで動画視聴行動や、学習効果にどのように違いが現れるかを、同一の講師による同一の喋り方で、編集方針だけ変えた 2 種類の動画を用いて実験を行った。実験協力者は高校生とし、積極的に参加をしてもらえるように、実際に学習している内容である英文法に関する動画とした。この実験を、実験タスクを柔軟に構成できること、実験者に負担をかけずに簡便に実験が実施できること、実験者の Personary にデータ集約できることを示す事例として挙げる。

学習者の動画視聴行動に注目した研究としては、動画配信型のオンライン学習サービスにおいて、動画プレイヤーの再生速度の初期設定を変更する介入を行なった場合に、1.25 倍速で再生した学習者の方が等速で再生した学習者よりも多くの修了証を得たという報告がある(Lang et al., 2020)。また、ある授業の講義動画に関して、同様に再生速度の初期設定に対して介入を行った場合、講義テーマに対する事前の習熟度テストの点数が低い学習者は再生速度 1.25 倍、事前の習熟度テストの点数が高い学習者は再生速度 1.5 倍の場合に、高い学習効果が得られることが示されている(Mo et al., 2022)。他に、講義動画の改善に注目した研究として、MOOC のログから講義動画とエンゲージメントの関係を分析した研究がある(Guo, Kim, & Rubin, 2014)。動画の長さで正規化した動画再生時間をエンゲージメントの指標として、動画配信型のオンライン学習サービスの再生記録を分析した結果、3 分以内の短い動画が高いエンゲージメントを持つと評価した。

講義動画の再生速度や長さに関する研究が多く、講義動画をどのような形式にするのが良いのか等、動画の質についての研究があまり見られないのは、比較実験のための動画を用意することが困難であるためと考えられる。すなわち、ある同じテーマに関する 2 つの動画教材があった場合、動画の長さ以外にも話者の外見的な特徴や話法、プレゼンテーション技術、撮影技法、編集技法といった多くの要素があり(Kurzweil et al., 2020)、既存の動画からこれらの要素が学習ログにどのような影響を与えているのかを推定することは困難である。また、実験としてこれらの要素を制御した動画を作成するのも、一定の指導スキルや編集スキルといった能力が要求されるため負担が大きい。

そこで我々は、東京大学開研究室の修士課程学生 I を実験者として、動画作成経験が豊富な予備校講師と協力し、同一の講義内容に関して、部分的に編集方針を変えた 2 種類の動画教材を使用した対照実験を実施する。本実験では介入する対象として、講義動画におけるチャプター分割およびその明示の有無に着目する。実験協力者である高校生にとって英語はすでに長期にわたって学習してきている内容である点に留意し、実験協力者の現在の理解度も考慮した分析を行なう。講義動画における

チャプター分割・明示の有無により、学習者の理解度ごとに、講義動画プレイヤーに対する学習者の行動がどのように変化するかを知ることで、学習効果の高い動画に関する示唆が得られる可能性がある。

3.2 実験方法

実験は学習者に動画視聴前にプレテストを実施してもらい、講義動画を視聴した直後にポストテストを実施してもらう形式とする。講義動画には2種類のタイプの動画を用意し、どちらの講義動画を視聴するかは学習者にランダムに割り当てるものとする。学習者に提示されるタスクの流れと講義動画のタイプを示す(図 13)。

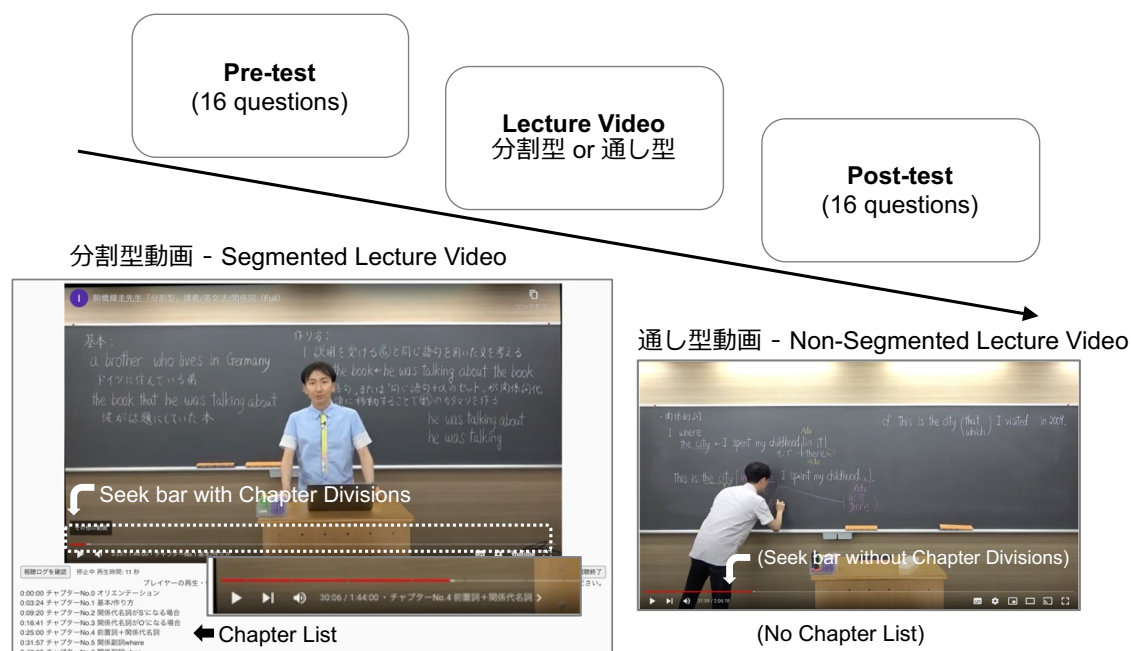


図 13 タスクの流れと用意した講義動画

本実験では英文法のうち 関係代名詞 をテーマとした講義動画をチャプター分割あり(分割型)・分割なし(通し型)として2種類を作成し、それらの動画に対する学習者の振る舞いを比較する。これら2つの動画はプロの予備校講師である駒橋氏に作成を依頼する。分割型動画は説明するトピックごとに動画をチャプター分割し、頭出しが容易なように、動画プレイヤーの画面の下にチャプター一覧を表示する。これに対して、通し型動画は講義の(単純な)録画映像のように、板書の様子も含めて動画としたものである。通し型動画にチャプター情報は表示されない。

動画はそれぞれ駒橋氏自身が講義を行うものとし、同じスタジオで実施し、撮影位置が同じになるようにする。また、理解度を確認するためのプレテスト・ポストテストの作成も、一般の高校生の英語の理解度を熟知する駒橋氏に依頼する。プレテストとポストテストは、4択式のものを2種類、駒橋氏に作成を依頼し作成する。プレおよびポストテストの内容はそれぞれ全学習者に対して同じものを使うものとし、

それぞれのテストは回答直後に学習者に対して問題ごとの正解・不正解と合計点数を表示するものとする。

3.2.1 講義動画の設計

講義動画は分割型、通し型いずれも 15 個のトピックから構成するものとする。トピック構成とその時間を示す(表 6)。

表 6 動画のトピック構成

No	Topic	通し (板書)	通し (板書以外)	分割
		Duration(sec)	Duration(sec)	Duration(sec)
0	オリエンテーション	0	50	204
1	基本/作り方	0	426	356
2	関係代名詞が S' になる場合	169	92	441
3	関係代名詞が O' になる場合	0	110	499
4	前置詞 + 関係代名詞	0	722	417
5	関係副詞 where	171	355	488
6	関係副詞 when	145	236	293
7	関係副詞 why	154	360	379
8	関係副詞が where, when, why が[名詞のカタマリ]をつくる場合	280	525	547
9	制限用法と非制限用法	162	696	529
10	,some of whom/,all of which など	234	380	384
11	whose + 名詞と the 名詞 of which	227	409	374
12	関係代名詞の what	179	298	526
13	関係代名詞 what を用いた慣用表現	85	637	366
14	いわゆる連鎖関係詞節	165	311	437
15	合計時間		5607	6240
		(板書含む 7578)		

分割型動画はこれらのトピックごとにチャプターとして分割され、そのトピックと開始時間の情報は動画プレイヤーのシークバー上と動画外の領域に一覧表示される。通し型は分割型のような明示的なチャプターはないものの、トピックの内容と説明順序は通し型と同様とする。それぞれ別に撮影したものであり、分割型は編集点を作りチャプター分割をしているが、通し型動画はひとつながりの分割がない動画となっている。分割型動画に比べ通し型動画の合計時間が長くなっているが、これは板書時間を含んでいるためである。なお、板書中は口頭による説明をしていない。また、特に分割型動画においてオリエンテーションにおいて、再生速度変更、シークを自由に実施してよい旨を説明している。

3.2.2 ログ収集システムの実装

動画プレイヤーは GO-E-MON の実験用タスクとして実装する。タスクスクリプトは GO-E-MON 開発者が動画プレイヤータスクを作成し、サンプルとして実験者 I に提供する。実験者 I は他の共同実験者と協力して、サンプルタスクに基づいて、プレテスト、講義動画再生、ポストテストから構成される実験タスクを実装する。一連の実験手順を単一の実験タスクとして実装することで、仮名 ID によってプレテスト、動画再生、ポストテストの点数変化及び行動データを追跡することができる。実験者 I はサーバー群の管理については意識することなく実験を実施することができる。

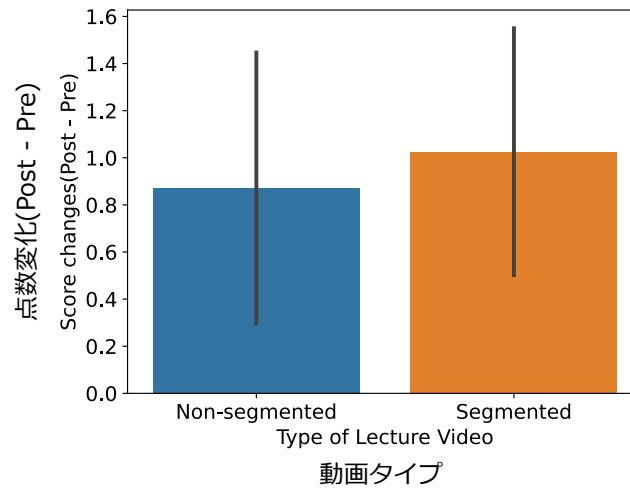
動画再生プレイヤーは実験 1 と同様に YouTube 組み込みプレイヤーを用いて実現するものとし、GO-E-MON 上で定義した JavaScript コードにより YouTube Iframe API(Google Developers, 2018)のイベントを取得しこれを動画視聴ログとして記録する。動画視聴ログには、プレイヤーの状態変更イベント(Ended, Playing, Paused, Buffering, Cued), 再生速度変更イベント, 再生品質変更イベントが記録される。これらのイベントの検出時に、その時点でのプレイヤーの再生状態(NotStarted, Ended, Playing, Paused, Buffering, Cued), 動画上の再生時刻, 再生速度, 再生品質を合わせて記録することで分析の助けとする。なお, YouTube Iframe API ではシークバーの操作イベントに対応する変化は記録できないため, 100msec ごとにプレイヤーの状態をポーリングし, 動画上の再生位置をチェックすることで, シークに関するイベントを検出する。スクリプトを自由に記述できる柔軟性を備えることで, API 等では得られない情報の収集を行うことができる。

3.3 実験結果

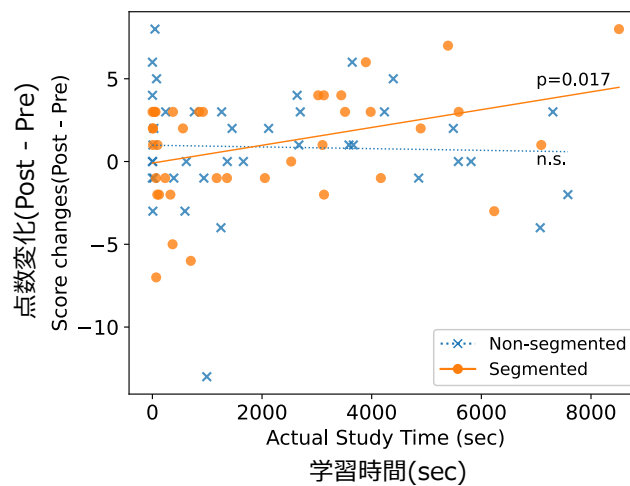
実験 2 の実験協力者募集は三浦学苑高等学校 1 年の進学コースに所属するクラスに対して行った。倫理審査 246-17 にしたがって実験内容を説明し、研究参加への募集を募った。研究参加への同意をし、かつポストテストまで実施した 78 名についてデータの分析を行なった。分割型と通し型は各生徒にランダムに割り当てた。78 名のうち分割型を実施したのは 39 名、通し型を実施したのは 39 名となった。研究参加に同意をしなくても生徒が不利益を被らないよう、学習アプリケーションは同意の有無に関わらず提供した。

実験者 I は他の共同実験者と協力して、GO-E-MON 開発者が提供するサンプルタスクに基づいて、プレテスト、講義動画再生、ポストテストから構成される実験タスクを実装することができた。実験者は作成した実験タスクの URL を実験参加者に配布することで、実験を実施できた。実験実施中、実験者は GO-E-MON のサーバー構成などの詳細を理解する必要なく、実験を遂行できることを確認することができた。ここからは、講義動画の構成方法に関して、本実験で得られた知見について紹介する。

分割型(Segmented)と通し型(Non-segmented)動画視聴後のプレテスト点数に対するポストテスト点数の増加量を示す(図 14).



(a)動画タイプ別 点数変化

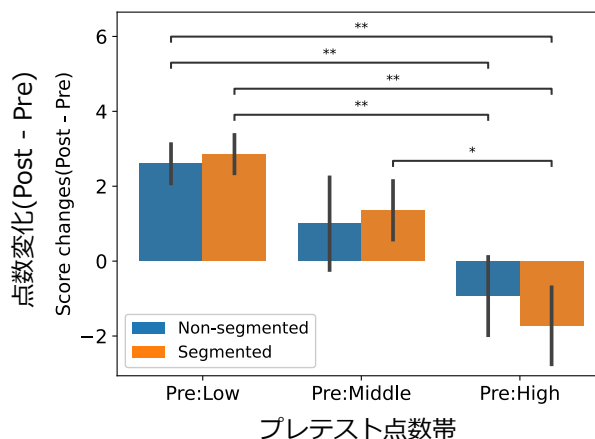


(b) 学習時間に対する点数変化

図 14 分割型と通し型における点数変化

動画タイプごとの点数変化(図 14(a))では平均では分割型が 1.03 点増加, 通し型が 0.87 点増加と, 分割型により成績向上傾向が見られるものの, t 検定による有意な違いは見られなかった. 一方, 実学習時間に対する点数向上(図 14(b))では分割型において有意な相関があった(点数増加と実時間データそれぞれの Q-Q 分布を確認したのち, ピアソンの相関係数で判定). 通し型は学習時間 1 時間程度までは点数向上の傾向は見られるものの, それ以上の長時間の学習に対して点数増加が見られなかった.

分割型, 通し型いずれもプレテスト点数が低い方が点数増加(Post - Pre)は高くなる傾向がある. プレテスト点数を 3 つの点数帯(Low, Middle, High)に分割し, 点数変化をまとめた(図 15).



プレテスト点数帯別の人数

Level	Pre-test(S)	Non-segmented	Segmented
Low	$S \leq 3$	15	14
Middle	$3 < S < 6$	9	14
High	$6 \leq S$	15	11

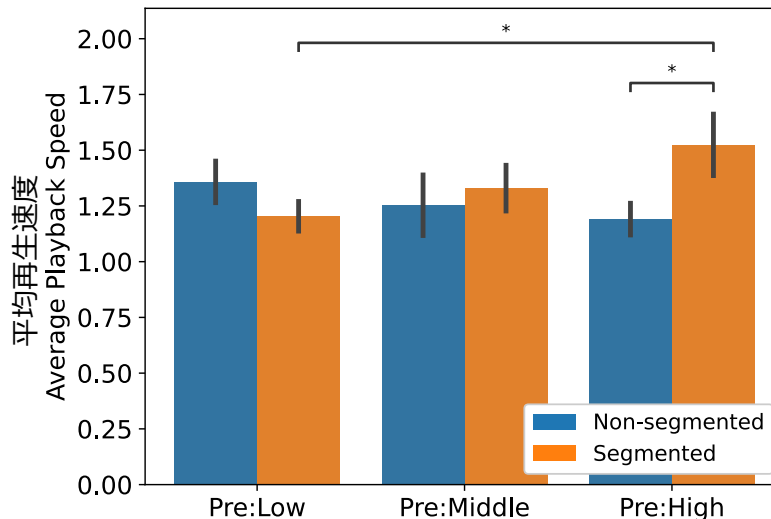
図 15 プレテストの点数帯による点数増加傾向の違い

*, **: t 検定により 2 つの値間に差がないとする帰無仮説が棄却されたケース. ** は $p < 0.01$, * は $p < 0.05$ を示す.

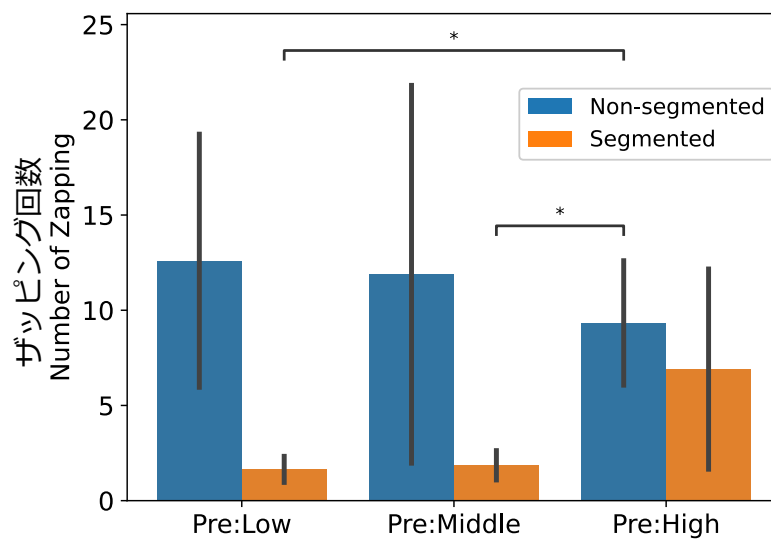
各点数帯の閾値は全員プレテストの点数を昇順にソートし, 参加者全体の 1/3 番目(小数点切り捨て)にあたる点数を Low の上限, 全体の 2/3 番目(小数点切り捨て)にあたる点数を High の下限とした. プレテストの点数帯別の点数変化を見るとプレテストが低い(Pre:Low)方が点数増加は多く, 高い(Pre:High)方が増加は少なくなる. 講義動画のうち基本的な部分は, プレテストが低い, すなわち英語が得意ではない学習者でもわかりやすいため点数増加に貢献したが, 応用部分は動画視聴だけでは十分に習得できなかった可能性が考えられる. また, High になると分割型通し型ともに平均はマイナス点となっている. このことから, プレ・ポストテストの難易度調整が十分ではなかった可能性が考えられる. プレ, ポストを交換しカウンターバランスを考慮することでより精度の高い点数増加に対する考察ができる可能性がある.

3.3.1 ビデオ視聴中の再生速度変更操作とザッピング操作

分割型動画と通し型動画で学習者による動画視聴中の操作がどのように違うのか、プレテスト点数帯ごとに平均動画再生速度とザッピング回数を示す(図 16)。



(a)プレテスト点数帯別の平均再生速度



(b)プレテスト点数帯別のザッピング回数

図 16 プレテストの点数帯と平均再生速度, ザッピング回数の関係

*: t 検定により 2 つの値間に差がないとする帰無仮説が棄却されたケース。 * は $p < 0.05$ を示す。

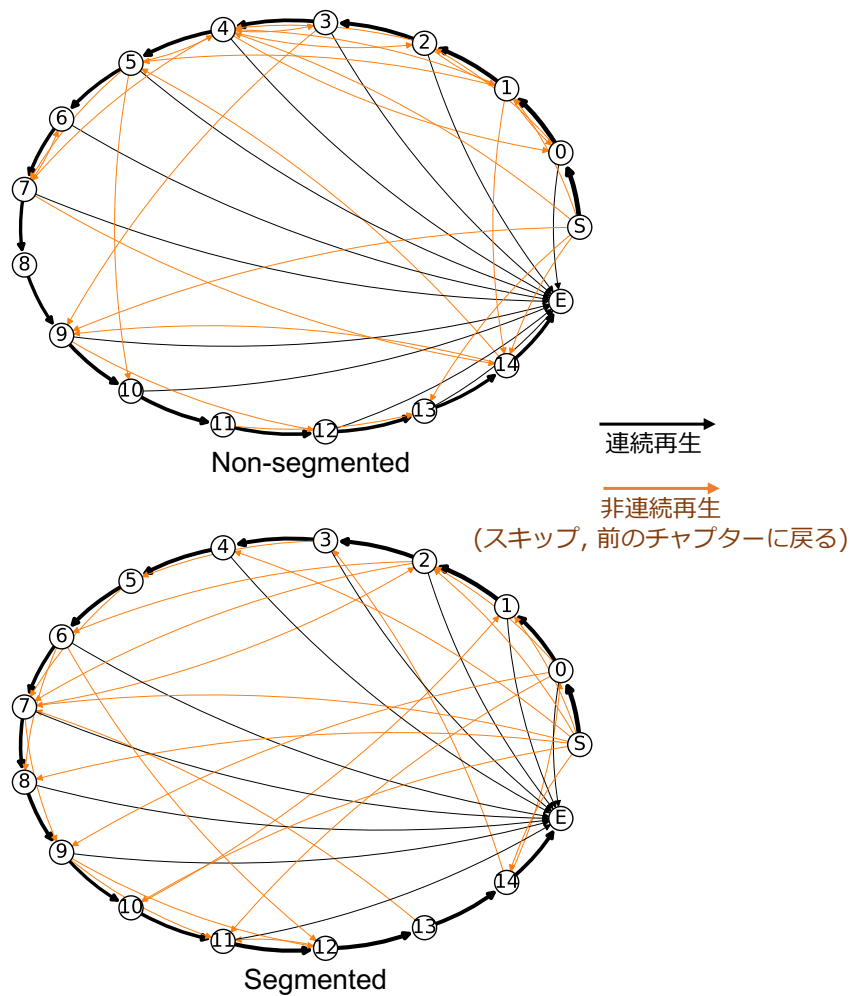
平均再生速度は連続視聴した(動画中の)合計時間をその再生に費やした実合計時間で割ったものである。また, ザッピング回数はシーク操作イベントのうち, 100msec 以上参照することをせず, すぐに別の再生位置に移動した回数を数えたものである。平均再生速度(図 16(a))に注目すると, 分割型動画を参照したプレテスト高成績群の学習者は再生時間を速めているのに対し, プレテスト低成績群の学習

者は実速度に近い速度で動画を参照している様子が確認できた。また、傾向としては有意ではないものの、通し型動画ではプレテスト低成績群が再生速度を速く、高成績群は実速度に近い速度で再生している傾向もあった。

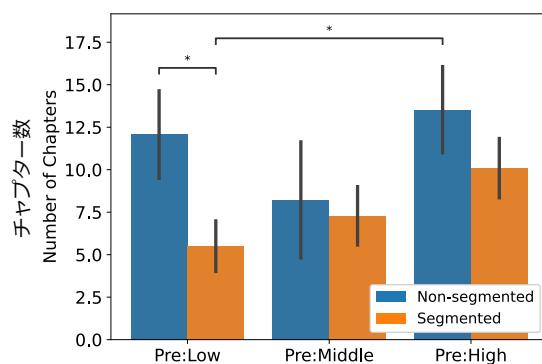
ザッピング回数(図 16(b))に注目すると、プレテスト点数が低い点数帯と中間点数帯に関して、分割型動画の方が通し型動画よりもザッピング回数が顕著に少ない傾向にあった。したがってプレテスト低～中得点帯の多くの学習者は通し型の学習者に比べて頻繁にシークバーを操作することなく、視聴時間を効果的に利用できた可能性が考えられる。このように、学習者はプレテストの点数帯と再生している動画のチャプターの有無によって異なる振る舞いを見せることを明確にできた。

3.3.2 再生箇所の選択

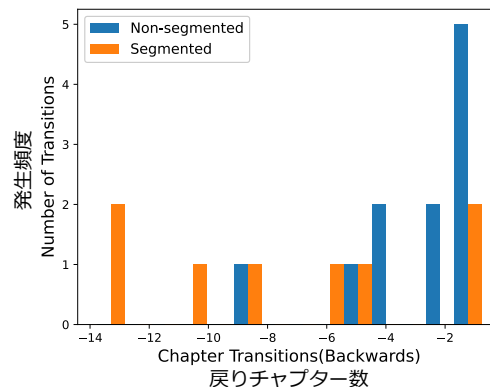
動画視聴ログから、動画中のどの部分がどのような順序で参照されているかを可視化することができる。トピックがどのような順序、分量で参照されているかを状態遷移図として示す(図 17)。



(a)講義動画タイプごとの状態遷移



(b)講義動画タイプ別 参照チャプター数



(c)講義タイプ別 前に戻る遷移の発生回数

図 17 プレテスト点数帯と参照チャプター数の関係

*: t 検定により 2 つの値間に差がないとする帰無仮説が棄却されたケース。 * は $p < 0.05$ を示す。

状態遷移(図 17(a))における状態名は S が再生開始, E が再生終了, 数字のものはトピック番号(15 秒以上参照したもの, 板書時間を含む)を示している. 矢印の太さは動画タイプごとに全ユーザーの遷移回数を合計したものである. どちらの動画タイプもオリエンテーションから順番に参照していくユーザーが多く, 途中のトピックで再生を止める学習者が多い. プレテストの点数帯ごとに学習者が視聴した合計トピック数の傾向(図 17(b))からは, 分割型の低得点層は通し型の低得点層や高得点層に比べて少ないチャプターしか参照していないことがわかる. また, 少数ではあるが前のトピックに戻って見返すような遷移(図 17(c))もあり, 通し型は 1 つ前のトピックに戻る遷移が多いのに対して分割型は大きく戻る遷移が見られる. チャプターが明示されることで, 学習者は短時間, 一部分でも確実に学習し, 長時間学習する学習者はわからなかった点を再度参照し復習できることが, 実学習時間に対して線形に効果が現れることに寄与している可能性がある.

3.4 考察

本実験では, GO-E-MON を用いて, 1 回の講義相当の短時間の実験を実施した. 実験者は講義動画の詳細な視聴ログを収集する実験タスクを, 提供されたサンプルコードを改変することで用意することができた. 実験の間, 実験者はシステム運用に対して意識する必要はなかった. また, 実験データは実験者の Personary に集約され, GO-E-MON に集中的に管理されることはない.

本実験では講義動画に関する知見も得ることができた. 講義動画視聴において学習者のプレテスト点数の高低, すなわち講義動画が対象とするトピックに関する理解度の高低と, 対象の講義動画の構成によって学習者が動画視聴行動を変えていることを示すことができた. 動画以外の学習行動から学習者の能動的・受動的な特性を検討している例があるが(Yoon, Lee, & Jo, 2021), 動画の提供スタイルを変える介入により, 学習者が自身の能力に合わせて適応的な動画視聴行動をおこなっていることを示した点で本実験はユニークである. 講義動画の作成作業は講義の実施経験や動画の編集スキルなど多くのノウハウを必要とし, 労力がかかる作業であるが, 本実験のアプローチは, 既存の動画の長さ, 再生速度のような定量的側面だけでなく, 構成方法などの質的側面も対照実験により評価可能になるという点で重要である.

3.4.1 仮名 ID の取り扱いにおける注意点

本実験において, 実験実施者の PDS に保存される実験結果ログは実験 1 と同様に全て仮名化された状態で保存されたが, 誤って実施校の教員によるタスク実施記録が混入する問題が発生した. 人数確認にてアカウント数に齟齬があったため, 実験実施者は GO-E-MON の開発・運用者に確認をした. この時, 生徒以外の Google アカウントによるアクセスが明らかになったため, この仮名 ID は分析時に除外した. オンライン実験はネットワークと計算機環境さえあれば参加でき, 実験室実験に比べて実施しやすい反面, 想定外のアクセスがあることを考慮する必要がある. これは実験 URL の配布時に関係者に対して実験協力者以外はアクセスしないようにすることなど, 正しく周知する必要がある. また今回は発生しなかったが, 実験参加同意の取り下げがあった場合にどのデータを分析から外すかを決定するためには, 実験協力者に確認するなど, 何らかの方法で仮名 ID と対象者の対応関係を取得する必

要が出てくる。仮名 ID と実名の対応づけは秘匿しつつ、データを適切に除外する機構が必要になると考えられる。

3.4.2 最適な動画再生速度に関する考察

本実験では分割型動画においてプレテスト点数が高いほど速度を速めて視聴する傾向を確認することができた。介入により適切な再生速度を探った研究では、習得度が低い群が 1.25 倍速、高い群は 1.5 倍速で良い学習効果が得られることが示されている(Mo et al., 2022)。本実験では、適切にチャプター分割がなされた動画であれば、学習者は外部から制御されなくとも、自身のレベルに合わせた再生速度調整を行なっている可能性を示すものである。通し型動画は分割型動画とは逆の振る舞いを学習者に行わせる可能性があるが、これは板書時間が影響している可能性がある。学習者は板書箇所の視聴中に実際の講義のように自身もノートに書き留めるといった行動をおこなっている可能性がある。このような場合、学習者が再生速度を落とす傾向が考えられる。対して、そのような行動をしない学習者にとって板書時間は認知的な負荷がかからない時間であり、結果的に再生速度変更を速める行動を選択したと考えることもできる。今後は、動画視聴中の動画再生以外の行動を収集することで、講義動画に学習者の適応的な行動の詳細を明らかにしていく必要がある。

3.4.3 最適な動画の長さに関する考察

複数の先行研究が、講義動画の最適な長さを数分程度と結論づけている(Guo, Kim, & Rubin, 2014; Zhu, et al., 2022)。短い動画において動画へのエンゲージメントが高まる、すなわちドロップアウトが少なくなることは、履修したかどうかを重視して成績評価をする学習サービスの場合は重要な知見となる。一方で本実験では、チャプター情報のような補助情報を与えることで、学習者が視聴箇所や視聴方法を取捨選択することができ、学習時間の有効活用につながっている可能性を示した。動画の途中で視聴を中断したとしても、中断までの学習時間に応じた学習効果は得られている。そのため、履修したかではなく、修得したかどうかを重視する学習サービスにとっては、適切なチャプター分割を行った、ある程度の長さを持つ動画を用意する方が、学習者に自身の理解度に応じた最適な動画視聴行動を促し、学習者が効率よく学習できる環境を提供できる可能性がある。

3.5 結論

GO-E-MON を使った実験の事例として、実験 2 について述べた。実験 2 では、実験者は GO-E-MON の実験タスクとして講義動画の視聴記録収集タスクを作成し、実験を行なった。実験者は、プレ・ポストテスト実施と講義動画プレイヤーでの動画再生中の行動収集をするタスクを、GO-E-MON 開発者が提供したサンプルに基づいて作成することができ、GO-E-MON の柔軟性を確認することができた。実験者は GO-E-MON 開発者が提供したサンプルに基づいて実験タスクの作成を行ったが、プレ・ポストテストを行うコードの追加は実験者自身が行なうことができおり、GO-E-MON 開発者以外でもタスクが作成可能であることを示す事例とすることができた。実験者はこのタスクを使用した実験を実施し、実験データを自身の Personary にて得て管理することができる安全性、また、実験者は実験に際してサ

ーバー管理などの技術的知識を要求されることはない簡便性についても実験 1 と同様に確認することができた。

実験 2 では、予備校講師の協力を受けることで、実験用に内容を調整した講義動画の作成を行った。Real World を指向した実験、特に長期に行う実験の場合には、参加者が実際に触れているものと同等の質の課題を提供しなければ、参加者が興味を維持することができず脱落してしまう危険性が高くなる。このようなコンテンツの作成にはコストがかかることが予想され、大きな課題であると考える。

4. 実験 3-1, 3-2: 英語学習コンテンツを用いた長期学習における学習動態の分析

特許出願予定の発明を含んでいるため、
第3章4.を削除し、関連する箇所を墨消しした
版です。

5. まとめ

本章ではオンライン実験プラットフォーム GO-E-MON を用いた複数の実験について、実験内容とその結果を報告した。

まず、GO-E-MON と分散型 PDS(Personary)の組み合わせが利用者にとって操作可能であり、実験データの安全性と流通性が得られているかを確認するため、大学生の PBL 形式の講義で GO-E-MON を使用する実験を行った(実験 1)。この実験により、実験者の Personary にて従来実験のようなデータ集約と分析が可能であること、実験協力者の Personary を用いた個人のデータ集約と再利用が可能であることを示した。GO-E-MON に長期のデータ管理は行わず、データを分散することで安全性を、実験協力者によるデータ管理を可能にすることでデータの流通性を実現する。また、実験者はサーバー等の存在を意識する必要がなく、実験実施自体は問題なく行えたものの、実験タスクの作成に対しては GO-E-MON 開発者がサンプルタスクを提供し、この設定部分を書き換えて利用する形となった。実験タスクを誰でも用意できるというわけではなく、簡便性に課題が残った。

残りは、GO-E-MON が可能なことや課題を示すための事例として、実際に GO-E-MON を用いて行った学習研究に関する実験について述べたものである。GO-E-MON の柔軟性と簡便性を活かして、短期的な(1 回の講義相当の)学習教材に対する学習者の行動データを収集する実験として、自学自習の場面における、講義動画プレイヤーに対する学習者の操作ログを収集する実験について紹介した(実験 2)。予備校講師に協力してもらい、同じトピックだが構成方針が異なる 2 つの動画を作成し、動画によって講義動画プレイヤーに対して学習者がどのような操作を行ったのか、詳細なログとして取得した。このログにより、学習者が自身の能力に応じて、動画に対するシーク、再生速度変更等の行動を変化させている様子を明らかにすることができた。ログ取得は YouTube IFrame API で行ったが、独自スクリプトによりシーク操作の検出も行なった。実験タスクスクリプトの柔軟性によりこれらのログの収集を実現することができた。また、実験データを自身の Personary にて得て管理することができる安全性、実験者は実験に際してサーバー管理などの技術的知識を要求されることはない簡便性についても実験 1 と同様に確認することができた。実験では動画の用意に関してコストをかけたが、Real World を指向した実験、特に長期に行う実験の場合には、参加者が実際に触れているのと同様の質の課題を提供しなければ、参加者が興味を維持することができず脱落してしまう危険性が高くなることが予想され、Real World 認知科学の実施における課題である。

また、長期的な(数日～数ヶ月の)学習教材と学習者のインタラクションを計測する実験として、高校生と大学生の協力を得て、英語学習アプリケーションに対する長期にわたる行動データを収集する実験について紹介した(実験 3-1, 実験 3-2)。実験 3-1 は 1000 人規模の参加者があり、GO-E-MON が多くの接続要求に耐える十分な可用性があることを示すことができた。また、これらの実験では、英語の学習教材における接触単語を、別課題での出題に利用する「随伴的出題法」について、長期にわたる学習動態の調査を行った。

最後に実験 3-2 ではウェアラブルセンサーを装着してもらうことで、生理指標データの同時収集も行い、GO-E-MON にて生理指標が同時収集できることとその可能性について検討を行った。ウェアラブルセンサーで得たデータを用いて、睡眠時間のばらつきと学習成績の間に従来研究と同様の傾向を確認することができた。本実験では随伴的出題に関して実験的介入を行ったのみで、生理指標に関連した介入は行っていない。GO-E-MON は生理指標の取得も容易にでき、生理指標に合わせて機械システムが応答を変化させるなどして、今後も知見が得られる可能性を示すことができた。生理指標を容易に安全に収集できることは非常に強力な一方、名寄せに利用されてしまう危険性がある。実験者の Personary に格納されるデータは仮名化されているが、たまたま同一の期間に生理指標を集めるような実験に実験協力者が参加していれば、名寄せに利用されてしまう危険性がある。そのため、仮名化よりも強力な名寄せ防止措置が必要である。

第4章 考察

1. Real World 認知科学実験への障壁

本研究では、Real World を指向した実験として、4 種類の実験について述べた。これらの実験は GO-E-MON 上でそれぞれ異なる実験者によって行われることで、実験結果データが実験者の PDS (Personary) にて分散管理された状況を実現することができた。また個々の実験者は Personary を連携させるための技術知識は要求されず、簡便に利用することができた。GO-E-MON は Real World 認知科学実験への障壁を緩和することができたと考えているが、まだいくつかの障壁が残っている。

1 つ目の障壁は実験タスクで利用するコンテンツの準備に関する障壁である。実験 2 では講義動画におけるチャプター分割の明示の有無がどのように学習者の振る舞いにどのような影響を与えるかの調査を行ったが、チャプターを分割したとしても、学習者がそれを妥当なものと感じることができなければ、学習者は利用しない可能性があった。本実験ではプロフェッショナルの講師に依頼することでこの問題を解決したが、これには労力がかかる。教科書等の既存のコンテンツを利用できれば、その負担は軽減できるが、既存の教育コンテンツは一般に著作権により保護されている。また、その著作権者もさまざまである。例えば教科書について考えた場合、出版社が著作権を保持する部分と記事等を記述した作家等が著作権を保持する部分が混在しており(日本図書教材協会, 2023)、これらのそれぞれの権利者に利用の許可を得ることは現実的ではない。

実験 3-1, 3-2 で用いた日本のマンガについては、出版社とマンガの作者が著作権を持っているため、実験目的を限定とした利用の許可を得ることができた。作者の死後 70 年経過により著作権が切れたコンテンツを教材として利用することで、この問題を回避することも考えられるが、古典とされるようなコンテンツに教材が偏ってしまうことで、学習者の関心を十分に引けないなど、期待する学習効果が得られないおそれもある。この制約は学習研究において教材を用いた介入がなされていない傾向があることの一因であると考えられる。学習科学に限らず、日常的に使用されているコンテンツを実験タスクとして用いる場合、改変に対して一定の制約があることは、Real World 認知科学の実現に向けての大きな障壁となる。

この問題の解決は容易ではないが、GO-E-MON が提供するシステム間連携機能が重要な役割を持つ可能性がある。教材側では Real World 実験に使われることを前提として、カスタマイズ可能な要素を用意しておき、実験 3-1 の出題単語の制御のように、URL パラメータ等で制御できるようにしておく。GO-E-MON 側での教材呼び出し時にこれらのパラメータを変化させることで、教材の応答を変化させることができる。ただし、このようなパラメータ化は教材コンテンツ開発者にとっては負担となるため、教材コンテンツ側にも恩恵がなければならない。Real World 認知科学に限らず、デジタル化が進む中で、学習の場に対して電子的にどのようにコンテンツを提供すべきかには多くの課題があり(井関, 2022)、今後の検討が必要である。

2 つ目の障壁は、実験協力者をタスクに継続的に参加させることの難しさである。実験 3-1 では 1000 名規模の大規模な募集を実施し実験を試みたものの、25%程度の実施にとどまった。これは実験

の考察で述べた通り、学習者の英語能力と実験タスクの難易度のミスマッチが原因の一つである可能性がある。先行研究(Okano et al., 2019)で示されている通り、適切なリマインドが必要であると考え、実験 3-2 で定期的なリマインドを実施し一定の効果を得たと思われるが、一方でリマインド頻度を過剰に増やしてしまうと逆効果になる可能性がある。GO-E-MON では比較的容易にリマインド処理を実装できるよう、メッセージングサービスとの連携機能と定期実行機能を実装したが、リマインドはただ送れば良いものではなく、動機付け研究(Higgins, 1998)を参考にしたり、実験開始後も継続的に実績を評価したりすることで改善を続けていく必要がある。そのため、Real World における実験を確実に進めるためには、実験終了後の事後分析だけではなく、実験中の参加状況などのデータ分析も重要になる。また、システムだけではなく、学習者と日常的に接している教員や関係者との連携も重要である。これは実験参加の有無が実験協力者の不利益にならないように配慮する必要もあり、実験協力への理解を得るための活動が実験室実験よりもより必要になる。

2. 実験結果データ管理における安全性のさらなる強化

本研究では、データの名寄せが容易にできないよう仮名化への配慮をしつつ、実験者と実験協力者の双方に実験結果データを送信するモデルを提案した。実験 1 の結果、オンライン実験環境 GO-E-MON を分散型 PDS と連携させることで、従来のように実験者が実験結果データを収集するだけでなく、実験協力者自身が実験結果データを活用できることを示すことができた。

仮名化がなされているとは言え、実験者が Personary アカウントに関して不適切な管理をすることで、実験結果データが流出する危険性は存在する。たとえ仮名化されていたとしても、実験 3-2 で得たように、1 週間の睡眠時間の変化を捉えた時系列データが流出してしまえば、その系列の値が一致するデータを検索し名寄せすることができてしまう可能性がある。そのため、実験者の Personary への実験結果データ送信に関して、仮名化よりも強力な変換処理を用意する必要がある。この変換処理の例として、匿名化処理がある。匿名化処理の例としては k-匿名化(Sweeney, 2002)が有名である。実験結果データに含まれる値の組み合わせの個数が最低 k 個となる、つまり、値から名寄せしようとしても必ず候補が k 個以上となる状態になるよう値を変換してから実験者の Personary アカウントにデータを送信することができれば、データが流出したとしても一定の安全性が確保される。

GO-E-MON と Personary との間の連携に関しては安全性向上の余地がある。GO-E-MON のアプリケーションサーバーは実験結果データを受信すると、実験者と実験協力者の Personary アカウントを識別して、適切なチャンネルにデータを書き込む。そのためアプリケーションサーバーに侵入された場合、Personary に送信される前の実験データが窃取される危険性がある。アプリケーションサーバーに一時的に保持される実験データの保持期間は、Personary が連携されてさえいれば、数分～数時間程度と短く、長期の実験データ漏洩のリスクは小さい。このリスクをさらに低減するためには、アプリケーションサーバーを複数の場所に分散させたり、アプリケーションサーバー内の実験データの保持期間をできるだけ短くしたりすることで、1 回の侵入で同時に窃取されうるデータを減らすことが可能である。なお、Personary アカウントが設定されていないアカウントのデータは、Personary が設定されるまで ResultStore の一時保存領域に保存される。GO-E-MON の実装では、データの漏洩リスクを考慮し、6 ヶ月経過後にデータを削除するものとした。実験結果データの機密性によっては時間をより短くするなどの選択肢を用意することで、より安全な管理が可能になる。

また、Personary のチャンネル共有方式における課題もある。GO-E-MON は Personary への書き込み時に、Personary に作成した GO-E-MON システムアカウントを使用する。実験者や実験協力者はこの GO-E-MON システムアカウントに対してチャンネルを共有することで GO-E-MON からの書き込みを受け入れるが、同時に GO-E-MON システムアカウントからのチャンネルの読み込みもできる状態になってしまう。実験者や実験協力者が実験終了後に共有を解除することでこの状態も解除されるものの、システムの Personary アカウントに対応する Google Drive アカウントおよび Personary アカウントのパスワードが窃取されると、現在共有されている GO-E-MON チャンネルを参照・編集できる危険性がある。このリスクを解消する方法として、Personary で「追記のみ」の共有設定を設けることが考えられる。実験協力者の Personary チャンネルを GO-E-MON に対して開示する際に、「追記のみ」で GO-E-MON に対して共有することで、読み込みを禁止することができれば、データ漏洩のリスクを大幅に軽減することができる。

「追記のみ」権限では GO-E-MON システムアカウントが窃取された場合に、大量のデータを書き込むことによるサービス拒否攻撃がされるおそれはあるが、これは適切なクォータなどの制限をかけることでリスクを軽減できる可能性がある。

3. 実験用スクリプトの記述の簡便性と安全性

本研究では実験用スクリプトを JavaScript により定義する形とし、GO-E-MON 開発者がいくつかのサンプルを作成して実験者に提供することで、各実験を実現した。GO-E-MON によりサーバー管理の手間は軽減するものの、スクリプトの作成は依然として障壁が高く、実験 1、実験 2 では GO-E-MON 開発者が提供したサンプルスクリプトを実験者に改変してもらうことでタスクを実現した。広く実験者に使ってもらうためには、このようなサンプルスクリプトによる助けなしに、実験者が行いたい実験を実現できる必要がある。JavaScript によって実験スクリプトを定義するフレームワークの代表的なものが jsPsych (de Leeuw, 2015)である。jsPsych を用いる場合でも、JavaScript でスクリプトを記述する必要があるものの、基本的な課題はプラグインの形で定義され配布されているため、JavaScript の記述における難易度は軽減される(黒木, 2020)。しかしながら、広く利用を促すためにはコードの記述作業自体を不要にすることが望ましい。このような要求に対しては lab.js (Henninger et al., 2021)を適用することが考えられる。lab.js はマウスによるグラフィカルな操作での質問フォームや課題の作成が可能なユーザーインターフェースを持つ。このような操作で作成可能なタスクは、lab.js があらかじめ用意しているものに限られるものの、lab.js を GO-E-MON に組み込むことでより実験タスク作成の障壁を下げられる可能性がある。

また、本研究ではユーザーが作成したタスクスクリプトに対して、Web ブラウザ上での実行についてはサンドボックス(Terrace, Beard, & Katta, 2012)などの実行制限を行わなかった。これは、実験者が記述したスクリプトに対して、デバイスやセンサーの利用を制限しないことを意図したものである。そのため、現状では暗号通貨マイニング(Huang et al., 2014)や不正なデータ取得(ReversingLabs, 2021)などの悪意のあるスクリプトが配布、実行される可能性がある。悪意あるコードの静的検出技術(Curtsinger, Zorn, & Seifert, 2011)やテスト環境(Musch et al., 2018)を用いることで、ある程度はこれらのコードを自動的に検出することが可能になる。また、実験者が自分の実験スクリプトを公開できる機能を実装することも有効である。スクリプトを公開することでより多くの研究者のチェックに晒されることになるが、その実験スクリプトを、別の実験者が自身の実験として行うことができれば、実験スクリプトを書くという形での研究への貢献も可能になる。

GO-E-MON では実験実施者が取得しているものと同じデータを実験協力者自身が取得することができるため、実験協力者自身が、実験結果データが実験説明から逸脱していないことを確認することが可能である。このようなチェックを確実にするためには、スクリプトが GO-E-MON 以外の方法で実験結果データを外部へと持ち出さないようにすることが重要である。また、実験データの形式を統一することで、実験データを積極的に実験協力者に開示しやすくなれば、実験への同意を得やすくなることも期待できる(Soldatova & King, 2006; Ocheja, Flanagan, & Ogata, 2018)。

4. 実験協力者自身の同意によるビッグデータ形成に向けて

GO-E-MON は分散型 PDS 技術を用いて、実験協力者のストレージに実験結果データを保存することで、実験協力者の同意による長期・縦断的なインタラクションデータの仮想的なプールの実現を目指す。研究者による実験データの二次利用を促進することで、より多くの知見を得ることができる可能性がある。実験協力者の同意によるデータ提供とそれによる仮想的な実験プールを健全な形で実現するためには、実験協力者が自身のデータの価値を理解することが重要である。しかし、実験 1 において、実験協力者にとってわかりにくいデータは、データ自体の価値が低く見積もられる可能性が示唆された。GO-E-MON は実験タスクの自由度を高く、簡便に書けるようにするため、Personary にデータを格納する形式は特に制約しないものとしたが、このことで Personary に蓄積されるデータ形式は多様なものとなり、データの価値を理解することも困難になることが考えられる。例えば本研究では実験の事例として学習行動に関するログを扱う実験を紹介したが、講義動画型の実験については YouTube の API のデータ構造に準拠した形式で実験結果データを保存し、学習アプリケーション型の実験については LTI Tool Provider が提供するデータ型に準拠した形式で実験結果データを保存した。

実際に他の研究者がこのデータを利用するためには、実験結果データ仕様を明らかにし、その仕様を広く公開する必要がある。このような仕様策定や文書作成は非常に手間がかかるため、適切な標準仕様があればそれに準拠するべきである。例えば学習ログに関しては Experience API(xAPI)がある(Advanced Distributed Learning Initiative, 2013)。xAPI においてログは「I did this(Actor-Verb-Object)」という単純な形式で記述することができ、Verb や Object の語彙は柔軟に拡張することが可能である。このような標準仕様に則ってデータを保持することで、実験協力者が実験実施者以外の研究者にデータ提供することが促進される可能性がある。具体的には、RDF(World Wide Web Consortium, 2001)などのメタデータを定義する機構を用いて、Dublin Core(The Dublin Core Metadata Initiative, 1995)のような共通の語彙を拡張した独自の語彙を定義することで、データ流通が促進される可能性がある。ただし、xAPI はデータの格納形式として JSON 形式(Ecma International, 2017)のファイルを想定するが、このようなファイルの処理は情報技術関連分野の研究者であれば容易だが、他分野の研究者にとってはそうではない。例えば Excel ファイルや MATLAB のような特定のアプリケーションの形式のファイルに慣れ親しんだ研究者も多いことが予測でき、標準化を考える場合はこれらのデータ形式への配慮も必要となると考えている。

実験協力者が自身のデータの価値を理解し、自身で管理していくようになるためには、自身の実験参加データを用いて自身が利益を得るといふ、一次利用的な体験も重要である。例えば、実験協力者が、自身の学習データを持ちそれを教育サービスに提供することができれば、実験 3-1 で得られた随伴的出題に関する知見をサービスとして利用することができる。

また、GO-E-MON で実施した実験ではないが、入眠前に学習する方が有利かどうかを検討した実験(Yazawa, Yoshimoto, & Hiraki, 2018)のデータを用いて、「あなたは入眠前に(あるいは日中に)勉強することでより良い成果が得られる可能性があります」等のアドバイスをするサービスを提供することも考えられる。

二次利用として縦断的な研究を促進するためには、誰が、どの実験に参加したのかを記録し、研究者に対して仲介を行うようなサービス(メディエータ)が必要である。例えば被験者募集システムとして大学等で利用されている Sona Systems(Sona Systems, n.d.)では、Sona 上で実施されたことのあるタスクについては、そのログを記録することができるため、以前に特定の実験に参加した参加者に絞ってリクルートすることができる。GO-E-MON 自体は実験協力者のリクルートやデータ紹介の機能を提供することは想定していないが、実験協力者からそれまでの実験参加記録を開示してもらい、利用すべきサービスの提案や、参加すべき研究の提案を行うサービスを開発することで、研究における二次利用を促進できる可能性がある。

また、実験協力者が自身の実験データを確認できるようにすることは好ましくないとされる場合もある。例えばウェクスラー児童用知能検査第 4 版(WISC-IV)では、保護者への説明責任という観点で適切な結果説明が必要であるとしながらも、記入済みの記録用紙を保護者に戻したり専門的所見を伝えたりといったことはすべきではないとしている(日本文化科学社, 2012)。その理由として、記入済みの記録用紙を戻すことで実験協力者が予期しないトレーニングをしてしまい、質問紙としての精度が損なわれる点が挙げられている。また、WISC-IV に限らず、保育施設の申込時の適性検査に関して、事前に適性検査の訓練を行ってしまう(Marks, 2005)といったような問題もある。しかしながら、これらの検査結果を実験協力者が長期に所有することで、実験協力者が得られるサービスの可能性も否定できない。現在の Personary では自身が所有するデータはすべてその内容を確認することができてしまう。このような問題に対しては、実験者や検査実施機関などの公開鍵で実験結果データを暗号化することで、実施者しか復号できないような形式でデータを実験協力者にフィードバックすることが考えられる。

実験者以外の人間が、その実験の成果公開前に実験データを確認できるようになることの問題もある。実験計画自体はプレレジのような形で登録することが推奨されているが(長谷川ら, 2021)、実験データそのものを実験成果の公表前に確認可能にすべきかどうかは検討が必要である。この問題に対しても、実験実施直後は Personary に実験者の鍵にて暗号化した状態でデータを保存し、研究成果の発表などに合わせて復号可能になるようなデータ管理をすることが考えられる。このように、実験者のみが復号可能となるような暗号化を活用することで、「(その時点では)見せてはいけない」データも実験協力者に渡すことができる可能性がある。このような機構を分散型 PDS として Personary が提供すべきか、オンライン実験環境として GO-E-MON が提供すべきかといった点に関する検討は今後の課題である。

GO-E-MON は分散型 PDS である Personary に対して実験データの送信をすることで実験データの安全な管理を目指す。これは集中管理型のサービスのよう、サービスのデータベースを攻撃されることで全データが窃取されるようなリスクが少ない一方で、データ管理はそれぞれの Personary の所有者に委ねられることになる。そのため、乳幼児のような明らかにデータ管理ができない主体のデータをどのように扱うかという問題が生じる。Personary は Google Drive などのパブリッククラウドストレージを暗号

化データの保存先として利用する。Google Drive を利用するために必要な Google アカウントの作成可能な年齢は 13 歳以上であるが、Google アカウントにはファミリーリンク(Google LLC, 2017)機能があり、保護者の Google アカウントによる管理のもと年齢の下限なしに子どもの Google アカウントを作成することが可能である。子どもの Google アカウントに対しては、紐づけられた保護者の Google アカウントが、アクセス可能なコンテンツの制御などを行うことができる。現在のところ Personary はファミリーリンクには対応していないものの、乳幼児の Personary アカウントをあらかじめ作成し、保護者の管理下とすることで、GO-E-MON で得られた実験データを格納可能になる可能性がある。

自身あるいは子どものデータを実験実施者に提供するかどうかの適切な判断ができるよう、第三者的な情報提供・支援を行う機関も必要である。学生のデータから内定辞退を予測したり早期退職を予測したりと、実験協力者の不利益に繋がる可能性があるデータ利用について多くの批判が集まっている(福島, 2019)。各機関の研究倫理委員会でこのような研究は実施できないよう適切にチェックされるべきではあるものの、データ所有者がこのような不利益を受けうるデータ利用に対してはデータ提供しない、あるいは部分的なデータ提供にとどめるといった判断ができるよう、実験データに関する検証や情報提供を行うサービスが必要である。Real World 認知科学が健全に進められるよう、実験協力者によるデータ利活用の問題は十分に検討していく必要がある。

第 5 章 まとめ

1. 結論

本研究では Real World 認知科学の実現に焦点を当て、長期的、継続的にインタラクション研究を実施可能なオンライン実験プラットフォーム GO-E-MON を提案した。

本研究では GO-E-MON が備えるべき特性として、特に以下の 5 点を挙げた。

- (1) 実験課題の柔軟性: 実験者は自身で考案した実験タスクを定義することができ、タスクに関わる操作データやウェアラブルセンサーによる生理指標を収集できる。
- (2) データの安全性: 実験協力者のパーソナルデータである、長期にわたる行動データや生理指標を安全に管理できる。
- (3) 実験環境の簡便性: 実験者はシステム開発やサーバー運用の負担を負わずに実験を実現できる。
- (4) 実験環境の可用性: 長期的・継続的な実験のため、多くの利用者からのアクセスや、機器故障などのトラブルがあっても実験環境を提供し続けられる。
- (5) データの流通性: 実験協力者の同意のもと、実験データを別の実験者が統合することができ、多くの協力者の長期にわたるデータを縦断的に分析することができる。

柔軟性については、Web ブラウザで動作する実験スクリプトを配信可能にすることで実現した。また簡便性はこれらのスクリプトを配信するサーバー群をサービスとして提供することで実現した。実験タスクの配信機能以外にも外部サービスとの連携機能は、全てサービスとして提供することで、実験者はこれらの開発や管理の負担を負わなくて済む。これらのサーバー群はクラウドサービスを用いて構築することで、可用性を実現した。

GO-E-MON が実験者のサーバー準備の負担を肩代わりすることの欠点は、これらのサーバー群が集中管理されたものとなることである。このようなサーバーは侵入や攻撃の対象となるため、外部からの攻撃から適切に防御をしなければならない。そこで安全性については分散型 PDS 技術、特に橋田による Personary(Hasida, 2014)を用いた。実験結果データを GO-E-MON のサーバー群が長期に集中管理するのではなく、実験実施後可能な限り速やかに実験者及び実験協力者の PDS(Personary)に実験結果データを送信する。送信後、サーバー群からはデータを削除することで、実験結果データが集中管理されることを防ぎ、安全性を向上する。GO-E-MON は、実験者が自分の実験結果データと他の実験者のそれとを容易に統合できないようにするため、メールアドレスやユーザ ID などの情報を付加することはせず、その代替として実験タスク及び実験協力者ごとに異なる仮名 ID を生成する。この実験データの分散管理の仕組みにより、アプリケーションサーバへの侵入による大規模なデータ漏洩を防ぐとともに、万が一、データが漏洩しても実験実施者のデータが他のデータと容易に突合されにくくなるよう設計した。また、実験協力者の Personary にデータをフィードバックすることで、実験協力者の意思によ

る実験データの再利用が可能になり、実験協力者の同意による仮想的な実験プールの実現も期待できる。

本研究では GO-E-MON と Personary が Real World 認知科学のためのオンライン実験環境として必要な特性を持っていることの検証を行うため、実験 1、実験 2、実験 3-1、実験 3-2 という 4 つの実験を事例として示した。実験 1 は GO-E-MON と Personary が提供する機能を各者が操作できることを大学生の PBL 形式の講義により調査した実験である。この実験により、実験データは実験者の Personary にて分散管理でき、安全性が確保可能であること、実験協力者の Personary を用いたデータの利用により、データの流通性が実現できることを示した。

次に、GO-E-MON の柔軟性と簡便性を活かした、Real World 認知科学を指向したインタラクション実験の事例として実験 2 について述べた。実験 2 は、短期間(数時間程度)のインタラクションの例として、高校生を対象に講義動画プレイヤーに対する学習者の行動を調査する実験である。予備校講師に協力してもらい、同じトピックだが構成方針が異なる 2 つの動画を作成、対照実験を実施し、動画の違いによって動画プレイヤーに対して学習者の振る舞いがどのように変化するか、詳細なログの分析を行った。実験結果から、講義動画にチャプター分割が明示されていることによって、学習者が自身の能力に応じて動画プレイヤーに対する再生速度変更等の行動を変化させている様子を明らかにすることができた。動画プレイヤーを表示して、その操作ログを収集するという独自の実験タスクを記述することができる柔軟性を示すことができた。

次に実験 3-1、3-2 を、可用性を要求される長期(数日～数ヶ月)にわたる実験の事例として示した。

最後に、ウェアラブルセンサーを用いた生理指標収集の事例として実験 3-2 について述べた。実験 3-2 では実験 3-1 の課題に加えてウェアラブルセンサーを装着してもらい、メッセージ送信機能を用いたリマインド等も組み合わせることで継続的なデータ取得の試行を行なった。ウェアラブルセンサーで得た生理指標データを用いて、学習成績と睡眠時間のばらつきとの間に従来研究と同様の傾向を確認することができた。センサーやメッセージングサービスを容易に組み合わせることで、従来にはなかったさまざまな実験を、サーバー管理等に慣れていない実験者も実施できる。

本研究では、Real World 認知科学を指向したオンライン実験環境を構築し、Real World 認知科学の特徴である、実世界でのインタラクションの動態を収集する実験を実施し、この実験から新たな知見を得られることを示すことができた。GO-E-MON は、高い柔軟性と簡便性を備えるため、従来は技術的障壁などでアイデアはあったが実現できていなかった実験を実施できる可能性がある。またデータの安全性と流通性によって、従来にはできない期間のデータ取得や大規模な分析ができる可能性がある。GO-

E-MON を用いて今後も Real World 実験を行うことで、今後も新たな知見が得られる可能性を示すことができた。

2. 今後の課題

GO-E-MON は多くの実験者に、広く、Real World 認知科学を指向した実験プラットフォームとして利用されるよう、今後いくつかの課題に取り組む必要がある。特に実験の事例から得られた今後の課題としては、以下のようなものが挙げられる。

(1) 日常的な課題を実験に利用できる柔軟性: GO-E-MON はさまざまな形態の実験を可能にする柔軟性を有するが、その実験で用いる刺激等のコンテンツの準備には課題が残る。例えば実験 2、実験 3-1 では教材が重要な役割を担ったが、一般に流通している教材の多くは、著作権により保護されているため、配信や改変を行うことが困難である。作成者の権利を適切に守りながら、実験として利用できるような枠組み、連携機能を整備していくことが重要である。

(2) 実験者に対するより安全なデータ共有形式の検討: GO-E-MON は分散型 PDS によりデータの安全な共有を実現するが、実験者へのデータ集約への配慮は仮名化では不十分である可能性がある。例えば実験 3-2 で得られた生理指標データは、同一期間に別の実験者が同様のデータを収集していれば、名寄せに利用されるおそれがある。実験者によるデータ取得モードに関して、統計情報のみの取得やデータの匿名化など、実験データから個人情報を排除した形でデータを取得することで名寄せや漏洩のリスクをさらに下げることができる。

(3) 柔軟性を維持しながらの実験タスクの簡便性、安全性向上: GO-E-MON はサーバー群をサービスとして提供することで、実験者がシステムの管理をする手間を不要にすることができたが、スクリプトを記述することは実験者によっては簡単ではない作業である。実験 1、実験 2 ではサンプルスクリプトを GO-E-MON 開発者が提供し、実験者がこれを変更することで実験を行ったが、ビジュアルエディタの導入など、スクリプトに対する苦手意識がある実験者でも利用しようと思える環境の導入が必要である。

(4) 実験協力者が実験データを所有することへのインセンティブ強化: GO-E-MON は分散型 PDS により実験協力者が自身の参加した実験データを所持し、開示可能な枠組みを、情報システムとして実現することができた。しかし多くの実験協力者がこれらのデータを管理するようになるためには、実験協力者がデータの価値を感じられるようにする必要がある。実験 1 では、解釈がしづらいデータに対しては価値が少なく評価される可能性を見出すことができた。実際に実験協力者がこのシステムを使って自身のデータを管理しようと思えるようにするためには、二次利用的な研究としての価値をアピールしていくことだけでなく、一次利用的なサービスとしての価値も示していく必要がある。

今後は、これらの課題に取り組むことで、Real World 認知科学の実現のため GO-E-MON をより良いオンライン実験プラットフォームとして改善していくつもりである。

謝辞

本研究を進めるにあたり終始あたたかいご指導と激励を賜りました東京大学 開一夫教授に心から感謝の意を表します。また博士論文としてまとめるにあたり、東京大学 藤垣裕子教授、馬場雪乃准教授、橋田浩一教授、大阪大学 鹿子木康弘教授には多大なるご指導をいただきました。深く感謝いたします。

各実験では多くの方のご協力をいただきました。実験 1 では東京大学 坂口菊恵特任准教授(現 独立行政法人大学改革支援・学位授与機構 研究開発部 教授)、開研究室修士課程(当時) 青木健氏に教材準備、環境準備など多大なるご支援をいただきました。実験 2 では三浦学苑高等学校 野櫻慎二先生、里見美郎先生には高校生の皆様にご協力いただくにあたり様々なご支援をいただき、また、駒橋輝圭氏には講義動画作成のご協力をいただきました。実験 3-1 では教材アプリケーション開発を株式会社 WillBooster の皆様、株式会社トレンディの皆様にご協力いただき、短期間での教材アプリケーション開発へのご協力をいただきました。また、自由英作文課題の評価システムとして SCREW を慶應大学今井倫太教授にご提供いただきました。実験 3-2 では、明治大学 嶋田総太郎教授、明治大学修士課程 永井口直也氏に実験を実施いただきました。ここですべての方のお名前をあげることはできませんが、オンライン実験プラットフォーム GO-E-MON に対して、様々な研究者の方から有用なフィードバックをいただき、ここに形にすることができました。関わってくださったすべての皆様の助言・支援・協力・励ましに対し、深く感謝いたします。

最後に、これまで私をあたたく応援してくれた家族に心から感謝します。

谷澤 智史

参考文献

- 1EdTech Consortium. (2019). Learning Tools Interoperability Core Specification 1.3, 16 April 2019. Available online: <https://www.imsglobal.org/spec/liti/v1p3> (accessed on March 12, 2023).
- Adamczyk, P. D., & Bailey, B. P. (2004, April). If not now, when? The effects of interruption at different moments within task execution. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 271-278.
- Advanced Distributed Learning Initiative. (2013). Available online: xAPI-Spec, <https://github.com/adlnet/xAPI-Spec> (accessed June 20, 2023)
- Akter, M., Gani, A., Rahman, M. O., Hassan, M. M., Almogren, A., & Ahmad, S. (2018). Performance Analysis of Personal Cloud Storage Services for Mobile Multimedia Health Record Management. *IEEE Access*, 6, 52625-52638. doi: 10.1109/ACCESS.2018.2869848.
- Aliyar, M & Peters, E. (2022). Incidental Acquisition of Italian Words from Comic Books. *Reading in a Foreign Language*, 34. 349-377.
- Almuhimedi, H., Schaub, F., Sadeh, N., Adjerid, I., Acquisti, A., Gluck, J., Cranor, L.F., & Agarwal, Y. (2015). Your Location has been Shared 5,398 Times! A Field Study on Mobile App Privacy Nudging. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, pp. 787-796. <https://doi.org/10.1145/2702123.2702210>
- Amazon Mechanical Turk, Inc. (2005). Amazon Mechanical Turk. Available online: <https://www.mturk.com/> (accessed on August 16, 2021).
- Amazon Web Services, Inc. (2006). Cloud Computing Services, March. Available online: <https://aws.amazon.com/> (accessed on March 13, 2023).
- 安西 祐一郎 (2011) 心と脳—認知科学入門, 岩波新書.
- Apple Inc. (2014). HealthKit. Available online: <https://developer.apple.com/documentation/healthkit> (accessed on March 12, 2023).
- Apple Inc. (2023). Safari 16.4 Beta Release Notes, 7 Mar 2023. Available online: https://developer.apple.com/documentation/safari-release-notes/safari-16_4-release-notes (accessed on March 12, 2023).
- Baker, M. (2016). 1,500 Scientists lift the lid on reproducibility. *Nature*, 533(7604), 452-454. Doi:10.1038/533452a
- Birch, K., Cochrane, D., & Ward, C. (2021). Data as asset? The measurement, governance, and valuation of digital personal data by Big Tech. *Big Data & Society*, 8(1). <https://doi.org/10.1177/20539517211017308>

- Burns, M. K. (2002). Comprehensive System of Assessment to Intervention Using Curriculum-Based Assessments. *Intervention in School and Clinic*, 38(1), 8–13.
<https://doi.org/10.1177/10534512020380010201>
- Camerer, C.F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Nave, G., Nosek, B.A., Pfeiffer, T., Altmejd, A., Buttrick, N., Chan, T., Chen, Y., Forsell, E., Gampa, A., Heikensten, E., Hummer, L., Imai, T., Isaksson, S., Manfredi, D., Rose, J., Wagenmakers, E.-J., & Wu, H. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behavior*, 2(9), 637–644. <https://doi.org/10.1038/s41562-018-0399-z>
- Center for Open Science. (2013). OSF. Available online: <https://osf.io/> (accessed on August 16, 2021).
- Charlie Browne Company Inc. (2013). New General Service List(NGSL) Available online:
<https://www.newgeneralservicelist.com/new-general-service-list> (accessed June 21, 2023)
- Chen, X., Jindal, A., Ding, N., Hu, Y. C., Gupta, M., & Vannithamby, R. (2015). Smartphone Background Activities in the Wild: Origin, Energy Drain, and Optimization. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking* (pp. 40–52).
<https://doi.org/10.1145/2789168.2790107>
- Classi 株式会社 (2020) Classi の現在の稼働状況についてお詫びとご報告. Available online:
<https://www.nihongakuen.ed.jp/s/20200514.pdf> (accessed on March 12, 2023).
- Cohn, N., & Kutas, M. (2017). What’s your neural function, visual narrative conjunction? Grammar, meaning, and fluency in sequential image processing. *Cognitive Research: Principles and Implications*, 2, 27. <https://doi.org/10.1186/s41235-017-0064-5>
- Corbett, A.T., & Anderson, J.R. (1992). The LISP intelligent tutoring system: Research in skill acquisition. In *Computer Assisted Instruction and Intelligent Tutoring Systems: Establishing Communication and Collaboration*. Erlbaum, 73–110.
- Corbett, A. T., & Anderson, J. R. (1995). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modelling and User-Adapted Interaction*, 4(4), 253–278.
<https://doi.org/10.1007/BF01099821>
- Cragg, L., Kovacevic, N., McIntosh, A.R., Poulsen, C., Martinu, K., Leonard, G., & Paus, T. (2011). Maturation of EEG power spectra in early adolescence: A longitudinal study. *Developmental Science*, 14(5), 935–943. <https://doi.org/10.1111/j.1467-7687.2010.01031.x>
- Curtsinger, C., Zorn, B., & Seifert, C. (2011). ZOZZLE: Fast and precise in-browser JavaScript malware detection. In *Proceedings of the 20th USENIX Security Symposium (USENIX Security 11)*.
- Dalipi, F., Imran, A. S., & Kastrati, Z. (2018). MOOCS dropout prediction using machine learning techniques: Review and research challenges. In *2018 IEEE global engineering education conference (EDUCON)*, 1007–1014. <https://doi.org/10.1109/EDUCON.2018.8363340>.

- Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227-268.
- DeepL. (2018). DeepL API. Available online: <https://www.deepl.com/en/docs-api> (accessed June 21, 2023)
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 4171-4186.
- Doleman, B., Williams, J.P., & Lund, J. (2019). Why most published meta-analysis findings are false. *Techniques in Coloproctology*, 23(9), 925–928. <https://doi.org/10.1007/s10151-019-02020-y>
- Donoghue, T., Voytek, B., & Ellis, S.E. (2021). Teaching creative and practical data science at scale. *Journal of Statistics and Data Science Education*, 29(sup1), S27–S39. DOI: 10.1080/10691898.2020.1860725
- Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology*, 84(5), 795–805. <https://doi.org/10.1037/0021-9010.84.5.795>
- The Dublin Core Metadata Initiative. (1995). Dublin Core. Available online: <https://www.dublincore.org/specifications/dublin-core/> (accessed on March 14, 2023).
- Duolingo. (2023). Official Website. Available online: <https://www.duolingo.com/> (accessed June 27, 2023)
- Eagle, N., Macy, M., & Claxton, R. (2010). Network diversity and economic development. *Science*, 328(5981), 1029–1031.
- Ecma International. (1997). ECMAScript® specification suite. Available online: <https://www.ecma-international.org/> (accessed on July 27, 2021).
- Ecma International. (2017). ECMA-404 The JSON Data Interchange Syntax. Available online: <https://www.ecma-international.org/publications-and-standards/standards/ecma-404/> (accessed on March 14, 2023).
- Evenson, K.R., Goto, M.M. & Furberg, R.D. (2015). Systematic review of the validity and reliability of consumer-wearable activity trackers. *International Journal of Behavioral Nutrition and Physical Activity*, 12, 159. <https://doi.org/10.1186/s12966-015-0314-1>
- Explosion AI. (2015). spaCy. Available online: <https://spacy.io/> (accessed June 21, 2023)
- Fett, A. J., Velthorst, E., Reichenberg, A., Ruggero, C. J., Callahan, J. L., Fochtmann, L. J., Carlson, G. A., Perlman, G., Bromet, E. J., & Kotov, R. (2020). Long-term Changes in Cognitive Functioning in Individuals With Psychotic Disorders: Findings From the Suffolk County Mental Health Project. *JAMA psychiatry*, 77(4), 387–396. <https://doi.org/10.1001/jamapsychiatry.2019.3993>

- Fitbit LLC. (2023). Fitbit Web API Reference. Available online: <https://dev.fitbit.com/build/reference/web-api/> (accessed on March 12, 2023).
- 藤田 早苗・小林 哲生・服部 正嗣・納谷 太 (2022) 日本語母語話者にとっての英単語親密度の調査. *言語処理学会第27回年次大会 (NLP-2022)*, 浜松 (オンライン開催), D1-3.
- 福島 直央 (2019) 就活サイト「内定辞退予測」で揺れる“個人スコア社会”到来の法的問題を考える. *情報法制研究*, 2019年6巻, 102-103
- Gagné, N., & Franzen, L. (2023). How to Run Behavioural Experiments Online: Best Practice Suggestions for Cognitive Psychology and Neuroscience. *Swiss Psychology Open*, 3(1), p.1. DOI: <https://doi.org/10.5334/spo.34>
- Gais, S., Lucas, B., & Born, J. (2006). Sleep after learning aids memory recall. *Learning and Memory*, 13, 259–262. DOI: <http://dx.doi.org/10.1101/lm.132106>
- Garmin Japan. (2012). Garmin vivosmart® 4. Available online: <https://www.garmin.co.jp/products/intosports/vivosmart-4-black-r/> (accessed on July 27, 2021)
- Garmin Ltd. (2021). Garmin Health API. Available online: <http://developer.garmin.com/garmin-connect-api/overview/> (accessed on July 28, 2021).
- Gickling, E., & Thompson, V. (1985). A personal view of curriculum-based assessment. *Exceptional Children*, 52, 205–218.
- Google Developers. (2018). Youtube IFrame Player API. Available online: https://developers.google.com/youtube/iframe_api_reference (accessed on July 28, 2021)
- Google LLC. (2008). Cloud Computing Services. Available online: <https://cloud.google.com/> (accessed on March 13, 2023).
- Google LLC. (2012). Firebase. Available online: <https://firebase.google.com/> (accessed on March 13, 2023).
- Google LLC. (2016). Cloud Vision API. Available online: <https://cloud.google.com/vision> (accessed June 21, 2023)
- Google LLC. (2017). ファミリー リンク開示情報(13 歳またはお住まいの国の該当する年齢未満のお子様がいる保護者向け), Available online: <https://families.google.com/familylink/privacy/notice> (accessed on September 7, 2023)
- Guo, P.J., Kim, J., & Rubin, R. (2014). How video production affects student engagement: An empirical study of MOOC videos. In *Proceedings of the first ACM conference on Learning @ scale conference (L@S '14)*. Association for Computing Machinery, New York, NY, USA, 41–50.
- Hajek, T., Kopecek, M., Alda, M., Uher, R., & Höschl, C. (2013). Why negative meta-analyses may be false? *European Neuropsychopharmacology*, 23(10), 1307–1309. <https://doi.org/10.1016/j.euroneuro.2013.01.004>
- Harris, C.R., Millman, K.J., van der Walt, S.J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M.H., Brett, M.,

- Haldane, A., Fernández del Río, J., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., & Oliphant, T.E. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357–362. <https://doi.org/10.1038/s41586-020-2649-2>
- 長谷川 龍樹・多田 奏恵・米満 文哉・池田 鮎美・山田 祐樹・高橋 康介・近藤 洋史 (2021) 実証的研究の事前登録の現状と実践——OSF 事前登録チュートリアル. *心理学研究*, *92*, 3, 188-196.
- Hasida, K. (2014). Personal life repository as a distributed PDS and its dissemination strategy for healthcare services. In *Proceedings of the AAAI 2014 Symposium on Big Data Becomes Personal: Knowledge into Meaning*, CA, USA, pp. 10–11.
- Henninger, F., Shevchenko, Y., Mertens, U. K., Kieslich, P. J., & Hilbig, B. E. (2021). lab.js: A free, open, online study builder. *Behavior Research Methods*, 1-18.
- Higgins, E. T. (1998). Promotion and prevention: Regulatory focus as a motivational principle. In *Advances in experimental social psychology* (Vol. 30, pp. 1-46). Academic Press.
- 開 一夫・今井 倫太・嶋田 総太郎 (2019) Real World 認知科学: インタラクション研究の新たな地平. *人工知能*, *34* 巻, 2 号, p. 236-244. https://doi.org/10.11517/jjsai.34.2_236
- Hofmann, W., & Patel, P. V. (2015). SurveySignal: A Convenient Solution for Experience Sampling Research Using Participants' Own Smartphones. *Social Science Computer Review*, *33*(2), 235–253. <https://doi.org/10.1177/0894439314525117>
- Hohmann, M.R., Konieczny, L., Hackl, M., Wirth, B., Zaman, T., Enficiaud, R., Grosse-Wentrup, M., & Schölkopf, B. (2020). MYND: Unsupervised Evaluation of Novel BCI Control Strategies on Consumer Hardware. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology (UIST '20)*. Association for Computing Machinery, New York, NY, USA, 1071–1084. <https://doi.org/10.1145/3379337.3415844>
- Hsu, H.C.K., Wang, C.V., & Levesque-Bristol, C. (2019). Reexamining the impact of self-determination theory on learning outcomes in the online learning environment. *Education and Information Technologies*, *24*, 2159–2174. <https://doi.org/10.1007/s10639-019-09863-w>
- Huang, D.Y., Dharmdasani, H., Meiklejohn, S., Dave, V., Grier, C., McCoy, D., Savage, S., Weaver, N., Snoeren, A.C., & Levchenko, K. (2014). Bitcoin: Monetizing stolen cycles. In *Proceedings of the Network and Distributed System Security Symposium (NDSS)*, 2014.
- Impey, C., & Formanek, M. (2021). MOOCs and 100 days of COVID: Enrollment surges in massive open online astronomy classes during the coronavirus pandemic. *Social Sciences & Humanities Open*, *4*(1), 100177, ISSN 2590-2911. <https://doi.org/10.1016/j.ssaho.2021.100177>
- InteraXon, Inc. (2021). Muse 2: Brain sensing headband. Available online: <https://choosemuse.com/muse-2/> (accessed on July 27 2021)
- Internet Engineering Task Force. (2012). RFC6749 – The OAuth 2.0 Authorization Framework. Available online: <https://www.rfc-editor.org/rfc/rfc6749> (accessed on March 12, 2023).

- Ioannidis, J.P. (2016). Anticipating consequences of sharing raw data and code and of awarding badges for sharing. *Journal of Clinical Epidemiology*, 70, 258–260.
Doi:10.1016/j.jclinepi.2015.04.015
- 井関 貴博 (2023) 大学の授業で試みた電子教科書のバンドル型ライセンスモデル: 大学と出版社のエコシステムを目指して. *デジタルアーカイブ学会誌*, 7 巻, 2 号, e11-e17. https://doi.org/10.24506/jsda.7.2_e11
- 板倉陽一郎 (2023). EHDS と日本の医療データ保護法制への示唆. *情報法制レポート*, 4, 17-35.
- Johnson-Laird, P.N., & Savary, F. (1999). Illusory inferences: A novel class of erroneous deductions. *Cognition*, 71(3), 191–229. ISSN 0010-0277. [https://doi.org/10.1016/S0010-0277\(99\)00015-3](https://doi.org/10.1016/S0010-0277(99)00015-3)
- Jordan, K. (2015) MOOC completion rates. Data. Available online:
<http://www.katyjordan.com/MOOCproject.html> (accessed on August 15, 2021).
- Kannan, V., Warriem, J.M., Majumdar, R., & Ogata, H. (2022). Learning dialogs orchestrated with BookRoll: effects on engagement and learning in an undergraduate physics course. *Research and Practice in Technology Enhanced Learning*, 17, 28. <https://doi.org/10.1186/s41039-022-00203-0>
- Karsh, N., Eitam, B., Mark, I., & Higgins, E. T. (2016). Bootstrapping agency: How control-relevant information affects motivation. *Journal of Experimental Psychology: General*, 145(10), 1333–1350. (2016-39039-001). Doi: 10.1037/xge0000212.supp
- Khan Academy. (2023). Official Website. Available online: <https://www.khanacademy.org/> (accessed June 27, 2023)
- Kim, J., Guo, P.J., Seaton, D.T., Mitros, P., Gajos, K.Z., & Miller, R.C. (2014). Understanding in-video dropouts and interaction peaks in online lecture videos. In *Proceedings of the first ACM conference on Learning @ scale conference (L@S '14)*, New York, NY, USA, 31–40.
<http://dx.doi.org/10.1145/2556325.2566237>
- Koval, P., Hinton, J., Dozo, N., Gleeson, J., Alvarez, M., Harrison, A., Vu, D., Susanto, R., Jayaputera, G., & Sinnott, R. (2019). SEMA3: Smartphone Ecological Momentary Assessment, Version 3. [Computer software]. Retrieved from <http://www.sema3.com>
- Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., Kelley, K., Hamrick, J., Grout, J., Corlay, S., Ivanov, P., Avila, D., Abdalla, S., Willing, C., & Jupyter Development Team. (2016). Jupyter Notebooks – A publishing format for reproducible computational workflows. In *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, pp. 87–90.
- クジラ飛行機 (2021) English-Japanese Dictionary “ejdict-hand”. Available online:
<https://github.com/kujirahand/EJDict> (accessed on March 13, 2023)
- 黒木大一郎 (2020) ウェブ実験の長所と短所, およびプログラム作成に必要な知識. *基礎心理学研究*, 38, 250–257

- Kuroki, D. (2021). A new jsPsych plugin for psychophysics, providing accurate display duration and stimulus onset asynchrony. *Behavior Research Methods*, *53*, 301–310.
<https://doi.org/10.3758/s13428-020-01445-w>
- Kurzweil, D., Marcellas, K., Henry, B., & Meyer, E. (2020). Evidence-Based Guidelines for Recording Slide-Based Lectures. *Medical Science Educator*, *30*, 1611–1616. <https://doi.org/10.1007/s40670-020-01032-w>
- Lang, D., Chen, G., Mirzaei, K., & Paepcke, A. (2020). Is faster better? A study of video playback speed. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge (LAK '20)*. Association for Computing Machinery, New York, NY, USA, 260–269.
 DOI:<https://doi.org/10.1145/3375462.3375466>
- Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., & Van Alstyne, M. (2009). SOCIAL SCIENCE: Computational Social Science. *Science*, *323*(5915), 721–723
- Liang, J., Li, C., & Zheng, L. (2016). Machine learning application in MOOCs: Dropout prediction. In *Proceedings of the 11th International Conference on Computer Science & Education (ICCSE)*, 52–57. Doi 10.1109/ICCSE.2016.7581554
- LINE Corporation. (2011). LINE. Available online: <https://line.me/> (accessed on March 13, 2023).
- de Leeuw, J.R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, *47*(1), 1–12. Doi:10.3758/s13428-014-0458-y
- Marks, A. (2005). Cracking the Kindergarten Code. Available online: <https://nymag.com/nymetro/urban/education/features/15141/index3.html> (accessed on September 7, 2023).
- Masuda, T., Batgerel, B., & Senzaki, S. (2020). Culture and Attention: Future Directions to Expand Research Beyond the Geographical Regions of WEIRD Cultures. *Frontiers in Psychology*, *11*.
<https://doi.org/10.3389/fpsyg.2020.01394>
- Matsumori, S., Okuoka, K., Shibata, R., Inoue, M., Fukuchi, Y., & Imai, M. (2023). Mask and Cloze: Automatic Open Cloze Question Generation Using a Masked Language Model. *IEEE Access*, *11*, 9835–9850. doi: 10.1109/ACCESS.2023.3239005.
- 松沼 光泰 (2006) 英語自己効力感(ESE)尺度の作成. 早稲田大学大学院教育学研究科紀要 別冊 issue 14-1(2006), 89-97
- McKinney, W. (2011). Pandas: A Foundational Python Library for Data Analysis and Statistics. *Python in High Performance Scientific Computing*, 1–9.
- Mesbahi, M.R., Rahmani, A.M., & Hosseinzadeh, M. (2018). Reliability and high availability in cloud computing environments: a reference roadmap. *Human-Centric Computing and Information Sciences*, *8*, 20. <https://doi.org/10.1186/s13673-018-0143-8>

- Meta Platforms, Inc. (2013). React.js, 29 May 2013. Available online: <https://reactjs.org/> (accessed on March 13, 2023).
- Millman, K.J., Brett, M., Barnowski, R., & Poline, J.B. (2018). Teaching computational reproducibility for neuroimaging. *Frontiers in Neuroscience*, 12, 727.
- Mo, C.-Y., Wang, C., Dai, J., & Jin, P. (2022). Video playback speed influence on learning effect from the perspective of personalized adaptive learning: A study based on cognitive load theory. *Frontiers in Psychology*, 13, 839982.
- de Montjoye, Y.A., Hidalgo, C.A., Verleysen, M., & Blondel, V.D. (2013). Unique in the crowd: The privacy bounds of human mobility. *Scientific Reports*, 3, 1376.
- de Montjoye, Y.A., Shmueli, E., Wang, S.S., & Pentland, A.S. (2014). openPDS: Protecting the Privacy of Metadata through SafeAnswers. *PLOS ONE*, 9(7), e98790.
<https://doi.org/10.1371/journal.pone.0098790>
- 文部科学省 (2010) 高等学校学習指導要領解説 外国語編・英語編. 東京: 開隆堂.
- 文部科学省総合教育政策局 (2021) 文部科学省 CBT システム(MEXCBT:メクビット)について. Available online: https://www.mext.go.jp/a_menu/shotou/zyouhou/mext_00001.html (accessed on March 12, 2023).
- 文部科学省初等中等教育局修学支援・教材課 2021, 端末利活用状況等の実態調査(令和3年7月末時点)(確定値). Available Online: https://www.mext.go.jp/content/20211125_mxt_shuukyo01-000009827_001.pdf (accessed on 18 Feb 2023).
- Moodle HQ. (2002). Moodle, 20 Aug 2002. Available online: <https://moodle.org/> (accessed on March 12, 2023).
- Musch, M., Wressnegger, C., Johns, M., & Rieck, K. (2018). Web-based Cryptojacking in the wild. *Preprint*, Arxiv abs/1808.09474.
- Narayanan, A.; Shmatikov, V. Robust de-anonymization of large sparse datasets. Secur. Priv. IEEE Symp. IEEE 2008, 2008. SP 2008, 111–125
- Narayanan, A. & Shmatikov, V. (2008). Robust de-anonymization of large sparse datasets. In *Proceedings of the 2008 IEEE Symposium on Security and Privacy (SP 2008)*, pp. 111–125.
- National Institute of Information and Communications Technology (Japan). (2008). Japanese Wordnet. Available online: <https://bond-lab.github.io/wnja/index.en.html> (accessed June 21, 2023)
- 日本文化科学社 (2012) 日本版 WISC-V テクニカルレポート #2 実施・報告の使用者責任と所見の書き方, Available online: <https://www.nichibun.co.jp/use/technicalreport/wisc4.html> (accessed on September 7, 2023).
- 日本学術会議 (2020) 提言 Web 調査の有効な学術的活用を目指して, Available online: <https://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-24-t292-3.pdf> (accessed on November 9, 2023).
- 日本図書教材協会 (2023) 学校用教材と著作権, Available online: <https://nit.or.jp/copyrightissue/> (accessed on November 22, 2023).

- Ocheja, P., Flanagan, B., & Ogata, H. (2018). Connecting decentralized learning records: A blockchain based learning analytics platform. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge (LAK '18)*. Association for Computing Machinery, New York, NY, USA, 265–269. DOI:<https://doi.org/10.1145/3170358.3170365>
- Okano, K., Kaczmarzyk, J.R., Dave, N., Gabrieli, J.D., & Grossman, J.C. (2019). Sleep quality, duration, and consistency are associated with better academic performance in college students. *Npj Science of Learning*, 4, 16. <https://doi.org/10.1038/s41539-019-0055-z>
- Open JS Foundation. (2006). jQuery: The Write Less, Do More, Javascript Library. Available online: <http://jquery.com/> (accessed on July 28, 2021).
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716. Doi:10.1126/science.aac4716
- Oxford University Press. (2009). Oxford Dictionaries API. Available online: <https://developer.oxforddictionaries.com/> (accessed June 21, 2023)
- Paivio, A. (1986). Mental representations: A dual coding approach. *Oxford University Press*.
- Papoutsaki, A., Sangkloy, P., Laskey, J., Daskalova, N., Huang, J., & Hays, J. (2016). WebGazer: Scalable webcam eye tracking using user interactions. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*, 3839–3845.
- Peirce, J., Gray, J.R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J.K. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods*, 51, 195–203. Doi:10.3758/s13428-018-01193-y
- Phipps, R. & Merisotis, J. (1999). What's the difference? A review of contemporary research on the effectiveness of distance learning in higher education. Washington, DC: Institute for Higher Education Policy. Available online: <https://eric.ed.gov/?id=ED429524> (accessed on June 20, 2023).
- Poulsen, A.T., Kamronn, S., Dmochowski, J., Parra, L.C., & Hansen, L.K. (2017). EEG in the classroom: Synchronised neural recordings during video presentation. *Scientific Reports*, 7, 43916. <https://doi.org/10.1038/srep43916>
- Powale, P. (2021). YouTube Engineering and Developers Blog Variable speed playback on mobile. Available online: <https://youtube-eng.googleblog.com/2017/09/variable-speed-playback-on-mobile.html> (accessed on August 15, 2021)
- Project jupyter team. (2019). JupyterHub. Available online: <https://jupyterhub.readthedocs.io/> (accessed on July 28, 2021)
- Rasch, B., & Born, J. (2013). About sleep's role in memory. *Physiological Reviews*, 93(April 2013), 681–766. DOI: <http://dx.doi.org/10.1152/physrev.00032.2012>
- ReversingLabs. (2021). Groundhog Day: NPM package caught stealing browser passwords. Available online: <https://blog.secure.software/groundhog-day-npm-package-caught-stealing-browser-passwords> (accessed on July 28, 2021).

- Roy, D., Patel, R., DeCamp, P., Kubat, R., Fleischman, M., Roy, B., Mavridis, N., Tellex, S., Salata, A., Guinness, J., Levit, M. & Gorniak, P. (2006). The human speechome project. In *Symbol Grounding and Beyond (Lecture Notes in Computer Science, Vol. 4211)*, pp. 192-196.
- Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25(1), 54-67.
- Sakai Project. (2005). Sakai LMS. Available online: <https://www.sakailms.org/> (accessed on March 12, 2023).
- Sano, A., Phillips, A.J., Yu, A.Z., McHill, A.W., Taylor, S., Jaques, N., Czeisler, C.A., Klerman, E.B., & Picard, R.W. (2015). Recognizing academic performance, sleep quality, stress level, and mental health using personality traits, wearable sensors and mobile phones. In *2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, June 9-12, 1-6. DOI: <https://doi.org/10.1109/BSN.2015.7299420>
- Sareen, P. (2013). Cloud computing: types, architecture, applications, concerns, virtualization and role of it governance in cloud. *International Journal of Advanced Research in Computer Science and Software Engineering*, 3(3).
- Sawangjai, P., Hompoonsup, S., Leelaarporn, P., Kongwudhikunakorn, S., & Wilaiprasitporn, T. (2020). Consumer Grade EEG measuring sensors as research tools: A review. *IEEE Sensors Journal*, 20(8), 3996-4024. Doi: 10.1109/JSEN.2019.2962874
- Schwab, K., Marcus, A., Oyola, J., Hoffman, W., & Luzi, M. (2011). Personal data: The emergence of a new asset class. An Initiative of the World Economic Forum. Available online: https://www3.weforum.org/docs/WEF_ITTC_PersonalDataNewAsset_Report_2011.pdf (accessed on November 13, 2023)
- Shaked, U. (2021) EEG Explorer. Available online: <https://github.com/urish/eeg-explorer> (accessed on July 28, 2021)
- Shiffman, S., Stone, A.A., & Hufford, M.R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, 4, pp.1-32.
- Shinya, Y. (2021). Usb: Universal-scale object detection benchmark. *arXiv preprint*. arXiv:2103.14027.
- Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, 46(5), 30.
- Skinner, B.F. (1938). The behavior of organisms: An experimental analysis. New York: Appleton-Century.
- Soldatova, L.N., & King, R.D. (2006). An ontology of scientific experiments. *Journal of The Royal Society Interface*, 3(11), 795-803. doi:10.1098/rsif.2006.0134
- Sona Systems (n.d.). Sona Systems: Cloud-based Participant Management Software. Available online: <https://www.sona-systems.com/>

- Statistics Bureau, Ministry of Internal Affairs and Communications of Japan. (2020). Diffusion of the Internet Ministry of Internal affairs and communications in Japan, Statistical Handbook of Japan. Available online: <https://www.stat.go.jp/english/data/handbook/c0117.html> (accessed on July 27, 2021).
- Stegman, P., Crawford, C., & Gray, J. (2018). WebBCI: An electroencephalography toolkit built on modern web technologies. In *Lecture Notes in Computer Science*, pp. 212–221.
- Stewart, N., Chandler, J., & Paolacci, G. (2017). Crowdsourcing Samples in Cognitive Science. *Trends in Cognitive Sciences*, 21(10), 736–748. <https://doi.org/10.1016/j.tics.2017.06.007>
- Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, 44(1), 24–31.
- Streeter, M. (2015). Mixture Modeling of Individual Learning Curves. In *Proceedings of the International Conference on Educational Data Mining (EDM)*, June 26-29, 45-52.
- Sweeney, L. (2002). k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(5), 557–570.
- Taylor, W.L. (1953). “Cloze Procedure”: A New Tool for Measuring Readability. *Journalism & Mass Communication Quarterly*, 30, 415 – 433.
- Tepgeç, M. & Ifenthaler, D. (2022). Learning analytics based interventions: A systematic review of experimental studies. In *Proceedings of the 19th International Conference on Cognition and Exploratory Learning in Digital Age (CELDA 2022)*, Lisbon, Portugal, pp. 327-330.
- Terrace, J., Beard, S.R., & Katta, N.P. (2012). JavaScript in JavaScript (js.js): Sandboxing third-party scripts. In *Proceedings of the 3rd USENIX Conference on Web Application Development (WebApps’12)*, USA, 9.
- Trampe, D., Quoidbach, J., & Taquet, M. (2015). Emotions in everyday life. *PLOS ONE*, 10(12), e0145450. <https://doi.org/10.1371/journal.pone.0145450>
- Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, İ., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., & SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3), 261–272. <https://doi.org/10.1038/s41592-019-0686-2>
- Walkington, C., & Bernacki, M. L. (2018). Personalization of instruction: Design dimensions and implications for cognition. *The Journal of Experimental Education*, 86(1), 50–68. <https://doi.org/10.1080/00220973.2017.1380590>
- Wang, R., Harari, G., Hao, P., Zhou, X., & Campbell, A.T. (2015). SmartGPA: how smartphones can assess and predict academic performance of college students. In *Proceedings of the 2015 ACM*

- International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*, September 07-11, 295–306. DOI: <http://dx.doi.org/10.1145/2750858.2804251>
- Web Bluetooth Community Group. (2021). Web Bluetooth draft community group report, 21 July 2021. Available online: <https://webbluetoothcg.github.io/web-bluetooth/> (accessed on July 27, 2021).
- Weller, M. (2018). Twenty years of Edtech. *Educause Review Online*, 53(4), pp. 34-48.
- World Wide Web Consortium. (1996). Cascading Style Sheets Specification. Available online: <https://www.w3.org/Style/CSS/Overview.en.html> (accessed on July 27, 2021)
- World Wide Web Consortium. (2001). Resource Description Framework (RDF). Available online: <https://www.w3.org/RDF/> (accessed on March 14, 2023).
- World Wide Web Consortium. (2022). Push API W3C Working Draft, 30 June 2022. Available online: <https://www.w3.org/TR/push-api/> (accessed on March 12, 2023).
- Yamamoto, E., Matsuda, G., Nagata, K., Dan, N., & Hiraki, K. (2019). Subtle temporal delays of mothers' responses affect imitation learning in children: Mother-child interaction study. *Journal of Experimental Child Psychology*, Vol. 179, pp. 126-142.
- Yang, Y., Youyou, W., & Uzzi, B. (2020). Estimating the deep replicability of scientific findings using human and artificial intelligence. *Proceedings of the National Academy of Sciences of the United States of America*, 117(20), 10762–10768. <https://doi.org/10.1073/pnas.1909046117>
- Yazawa, S., Yoshimoto, H., & Hiraki, K. (2018). Learning with Wearable Devices reveals Learners' Best Time to Learn. In *Proceedings of the 2018 2nd International Conference on Education and E-Learning (ICEEL '18)*. Association for Computing Machinery, New York, NY, USA, 87–92. <https://doi.org/10.1145/3291078.3291097>
- Yetano, A., & Royo, S. (2017). Keeping Citizens Engaged: A Comparison Between Online and Offline Participants. *Administration & Society*, 49(3), 394-422. <https://doi.org/10.1177/0095399715581625>
- Yoon, M., Lee, J., & Jo, I.-H. (2021). Video learning analytics: Investigating behavioral patterns and learner clusters in video-based online learning. *The Internet and Higher Education*, 50.
- ZDNET. (2020). Google: G Suite now has 2 billion users. Available online: <https://www.zdnet.com/article/google-g-suite-now-has-2-billion-users/> (accessed on July 28, 2021).
- Zhu, J., Yuan, H., Zhang, Q., Huang, P.-H., Wang, Y., Duan, S., Lei, M., Lim, E.G., & Song, P. (2022). The impact of short videos on student performance in an online-flipped college engineering course. *Humanities and Social Sciences Communications*, 9, 327.
- Zolfaghari, B., Srivastava, G., Roy, S., Nemati, H.R., Afghah, F., Koshiba, T., Razi, A., Bibak, K., Mitra, P., & Rai, B.K. (2020). Content delivery networks: state of the art, trends, and future roadmap. *ACM Computing Surveys (CSUR)*, 53(2), 1-34.