

あアラウド法における発声が体験者に与える負荷の調査

安中 勇貴^{1,a)} 能丸 天志¹ 山岸 丈留¹ 渡邊 恵太^{2,b)}

概要：エンタテインメントを評価するためにはコンテンツ体験者の心の動きを評価する必要があるが、心の動きに着目する評価手法は定まっていない。川島らは体験者に「あ」のみで感情を表現させ、発声の大きさや抑揚から感情を観測するあアラウド法を提案した。高山らは、あアラウド法において体験者と分析者が利用しやすい発声方法について検討したが、それぞれがよいと思う発声は異なり、両者にとって最適な発声方法は見つかっていない。本研究では体験者の発声中の身体的負荷と認知的負荷を計測し、発声方法が体験者に与える影響について調査する。得られた結果から、あアラウド法においてより良い発声方法についても考察を行う。

1. はじめに

エンタテインメントシステムの評価において、感情変容の観測が重要である。感情の評価はアンケートを用いた事後の主観指標であることが多い [1][2]。この指標は、システムを体験後の総合的な体験評価は行えるが、体験中の時系列的な感情変容はわからない。また、心拍や発汗、皮膚コンダクタンス反応などの生理指標を用いて、感情を観測することがある。これらの手法は、体験者のリアルタイムの状態がわかるが、不快感などの特定の感情を観測するだけに留まる。

この問題に対して、川島ら [3] は体験者に「あ」のみを発声させ、声の大きさや抑揚から感情を観測する「あアラウド法」を提案した。川島らがあアラウド法が思考発話法よりも言語化負担と反応性が少ないこと、「あ」の発声と感情に関連があることを示唆し、エンタテインメントシステムの評価への利用可能性を示した。高山ら [4] はあアラウド法における複数の発声方法でビデオゲームを行い、あアラウド法において適した発声方法について調べた。その結果、体験者は感情が生じたときに「あ」を発声する都度発声が、分析者は常に「あ」を発声し続け、感情に応じて「あ」の変化で表現する耐久発声を使いやすいとし、発声方法の使いやすさが異なった。ただどちらの実験においても、参加者の負荷についてリッカートスケールで評価しており、あアラウド法における発声が体験者に与える影響に



図 1 あアラウド法は「あ」の発声を用いて体験者のリアルタイムの感情を観測する感情評価手法である。

については十分に検討されていない。

そこで本研究では、あアラウド法における発声が体験者に与える影響について、主観的な負荷と生体反応から調査する。実験ではあアラウド法を含めた複数の発声方法でビデオゲームをプレイし、発声が体験者に与える負荷について瞳孔径と皮膚電気活動 (EDA)、NASA-TLX を用いて評価する。これらの指標から発声方法ごとの参加者の負担を評価し、体験者が感情表現しやすく、タスクの妨げにならない発声方法について検討する。

2. 関連研究

2.1 エンタテインメントシステムの体験評価における需要と課題

ユーザーの体験評価はエンタテインメントや製品開発において重要である [5][6]。例えばビデオゲームの体験評価では、アンケートによる主観的評価や時系列の生理的反応や脳波を用いて、ユーザーの関わりやパフォーマンス、幸

¹ 明治大学大学院 先端数理科学研究科 先端メディアサイエンス専攻

² 明治大学 総合数理学部 先端メディアサイエンス学科

a) cs232044@meiji.ac.jp

b) watanabe@fms.meiji.ac.jp

福感や興奮などを調査する [7][8][9]。また作業中の音楽やゲームのBGMが体験に与える影響の調査では、没入感や緊張感、満足度などを調べるために、主観的評価に加え皮膚電気活動や顔面筋活動を用いることがある。[10][11]。その他にもダンスや視覚アートでは、パフォーマンス後の意識的、内省的な振り返りによる質的評価を用いて、運動学的・共感的反応や快楽、期待などを評価する [12]。しかし、こうしたユーザーの体験評価は、ユーザーの反応や周囲の状況、コンテキストを考慮する必要がある。そのため、実験手順や分析方法、評価基準の設定が難しい。[13]。また、分野や研究によって実験設計や評価方法が異なるため、幅広い応用が難しい。

2.2 既存の感情観測の分類と比較

感情観測は大きく分けて、行動指標、生理指標、主観指標に分けられる。行動指標は、反応時間から表情、視線、姿勢、しぐさ、声色など、非言語的に表出される行動全般を指す。表情は、顔の各部位の動きをカメラでとらえ、感情を識別できる [14]。発声による感情観測は、声のピッチの強弱から感情を識別できる [15]。これらの手法は、非侵襲で被験者の負担が少なくリアルタイムな分析が可能である。さらに、深層学習を活用することで、迅速で正確な評価が行える可能性がある [16][17]。また Damacio ら [18] は、客観的に観測可能な身体反応を示す情動と、言語化が必要な主観的経験を示す感情を区別している。表情や姿勢などの行動指標は、情動の観測に優れる一方で、意図的な表現による感情の観測が困難である。

生理指標による感情観測は、心拍変動、皮膚電位、脳波、筋電、体温など、様々なアプローチをとる。生体反応は、無意識の反応を捉えられ、改ざんが困難である利点がある。これにより、感情を客観的に評価できる。Kreibig らは、驚きやリラックスなどの一部の感情が、特定の生態反応と強く関連していることを明らかにした [19]。一方で、すべての感情を生体反応で観測することは難しく、感情の解釈が明確でない場合もあると主張した。動きを伴う体験評価では、正確な計測が困難になる欠点もある [20][21]。加えて、計測機器を用いた感情観測には専門的な機器と知識を必要とし、装着や分析に時間とコストを要する。

主観指標は、質問紙や発話などの言語を介した主観的な感情経験の事後報告である。感情観測における質問紙は多岐にわたり、特定の感情や感情のスペクトル全体の評価に用いる。例として、PANAS[22] や SAM[23]、DEQ[24] などがある。特に UES[25] は、インタラクティブな体験の評価に特化した質問紙である。発話による観測には、レトロスペクティブ法や思考発話法など [26] がある。しかし、質問紙やレトロスペクティブ法では実験後に評価を行うため、リアルタイムに感情を得られず、忘却や捏造などのリ

スクがある。思考発話法は、参加者にタスク中の思考を発声させることで、体験者の主観的感情をリアルタイムで観測が可能である。一方で、体験者にとって考えていることを言語化する負担は大きく、タスクへの集中度の低下や感情が十分に表現できない可能性がある。

あアラウド法は主観指標と行動指標を用いた観測手法である。そのため生体反応などの生理指標と比較して、より直接的に体験者の感情を評価できる可能性がある。また、発声のような行動指標より、意図的な表現に適している。また、リアルタイムでの発声を行うため、レトロスペクティブ法で懸念される忘却や捏造が起こりにくい。観測手法である思考発話法と比較すると、感嘆詞である「あ」の発声のみを用いることで言語化の負担を軽減し、主観的な認知負荷などを軽減できる可能性がある [27]。

2.3 リアルタイムに評価をする手法における負荷

体験者がタスクと同時並行で発話やデバイスによる入力を行う手法は、実験参加者に負荷を与え、実験タスクに影響を与える可能性がある。長谷川 [28] らは思考発話法において、発話量の個人差とワーキングメモリ容量の関係、思考発話が課題成績に与える影響を調査した。実験の結果、課題難易度が変化すると、ワーキングメモリの容量によって発話量が異なった。また、発話によって解答時間が遅くなることが示され、発話が課題遂行を妨害する可能性があることを示した。加えて、リアルタイムな感情アノテーション技術の評価する時にはこれらのことに注意して評価することがある。Tianyi ら [29] はモバイルデバイスにおけるリアルタイム感情アノテーション手法を提案し、NASA-TLX、瞳孔径、心拍、皮膚電位の4つを用いて使用者の作業負荷やストレス、タスク集中度などを調べた。

あアラウド法は体験者が発声を行う感情評価手法であり、発声が体験中のタスクに影響を与える可能性がある。そのため、本研究ではあアラウド法がタスクに与える影響について、体験者の主観的な認知負荷と生体反応から調査する。

3. 実験

本研究の目的はあアラウド法における発声方法が体験者に与える負荷を明らかにすることである。

3.1 実験参加者

実験参加者は、19歳から24歳の男性6名、女性2名の計8名である。

3.2 実験条件

本実験では4つのパターンで実験を行った。あアラウド法の発声方法については高山らの実験 [4] を参考にした。発声なし

発声することなく、通常通りタスクを行う。参加者には発声はせずに、通常通りプレイするように指示した。

感情発話

感情の変化が起きたときにそれがどういった感情かを述べる。参加者にはゲーム中に感情が変化したときに、それがどういった感情かを「嬉しい」や「イライラする」などの発話で表現するように指示した。またゲーム中に操作方法を言い続けるなど過度に認知プロセスや感情と関係ない言葉を発話することがないように伝えた。

都度発声（あアラウド法）

感情変化が起きたときのみ「あ」と発声する。参加者にはゲーム中に感情が変化したときに、それがどういった感情かを「あ」の発声で表現するように指示した。

耐久発声（あアラウド法）

感情変化がなくとも、息が続く限り「あ」と発声し続ける。参加者にはプレイ中は「あ」を発声し続け、感情の変化したときにそれを「あ」の抑揚や大きさの変化で表現するように指示した。

3.3 実験タスク

本実験は、飯田ら [30] の Open Video Game Library のビデオゲームの一つである Sliding Penguin を利用した。このゲームは氷の上をすべるペンギンを操作し、道中に落ちていいる魚を獲得しながら制限時間内にゴールを目指すゲームである。実験では制限時間が 45 秒、得点として道中に魚が 15 個配置したコースで実験を行った。実験参加者にはできるだけ多くの魚を獲得し、制限時間内にゴールするように指示した。

3.4 実験環境

実験参加者は左手首に E4 リストバンドを装着した。これは実験参加者の EDA を取得するためであり、4Hz で EDA を取得した。プレイには DELL の 27 インチ 4K モニターと Xbox コントローラを使用した。またディスプレイの下には瞳孔径を計測するために Tobii Pro Fusion を設置し、30Hz で駆動させた。また、実験でのアンケートに回答するために、iPhone15 を使用した。

3.5 実験手順

実験の流れを以下に示す。

- (1) 実験同意書に記入する
- (2) 実験説明を受ける
- (3) ゲームの練習を行う
- (4) それぞれの指標のベースラインを取得する
- (5) 指示された発声でゲームをプレイする
- (6) NASA-TLX に回答する

NASA-TLXにおけるWeight Ratingのスコア

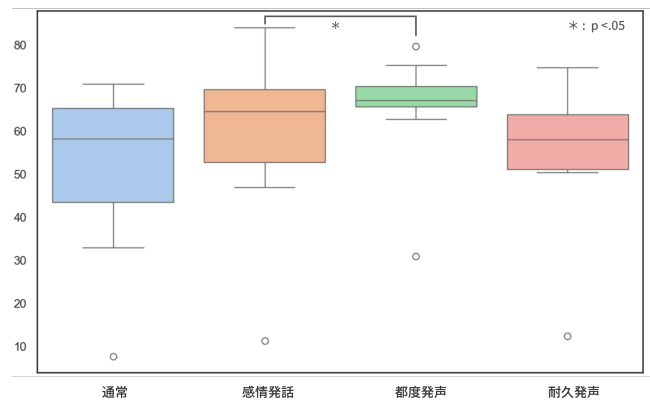


図 2 NASA-TLX における Weighted Rating のスコア

(7) (3) (4) をすべての発声方法で行う

ベースラインの取得は画面を黒一色にし、被験者に 10 秒間画面を見つめさせ、その状態で心拍と皮膚電位、瞳孔径を取得した。NASA-TLX は NASA が AppStore で配布しているアプリケーションを iPhone15 にインストールし、利用した。

3.6 結果

NASA-TLX、瞳孔径と EDA、半構造化インタビューから収集したデータを分析した。

NASA-TLX の結果を図 2 に示す。発声方法ごとに NASA-TLX の Weighted Rating の値を比較した。Shapiro-Wilk 検定の結果、データが正規分布でないことが分かったため、フリードマン検定を行ったところ有意差 ($p < 0.05$) があった。そこで Bonferroni 法で補正を行い wilcoxon の符号順位和検定で多重比較を行ったところ、通常と感情発話、都度発声と感情発話、耐久発声と感情発話の 3 つで有意差 ($p < 0.05$) があった。このことから、感情発声は通常時及びあアラウド法と比較して作業負荷が高いことが示された。

発声方法ごとに瞳孔径と EDA の変化を比較した。タスク前に取得した 10 秒間のデータの平均値をベースラインとして、それぞれの値の変化量を調べた。瞳孔径のデータを図 3 に示す。Shapiro-Wilk 検定の結果、データが正規分布であったため、分散分析を行ったが有意差は認められなかった。EDA のデータを図 4 に示す。Shapiro-Wilk 検定の結果、データが正規分布でないことが分かったため、フリードマン検定を行ったが有意差は認められなかった。

実験後に 3 つの発声方法ごとに感情表現ができたか、ゲームに集中できたかを半構造化インタビューを行った。感情表現に関するインタビューでは実験参加者 8 人のうち 6 人が耐久発声、2 人が都度発声最も感情表現しやすかったと回答した。耐久発声を選んだ人からは、常に発声をしているので感情を発声で表現するためのハードルが低く、何気ない小さな感情を表現で切ること、反射で反応がしやす

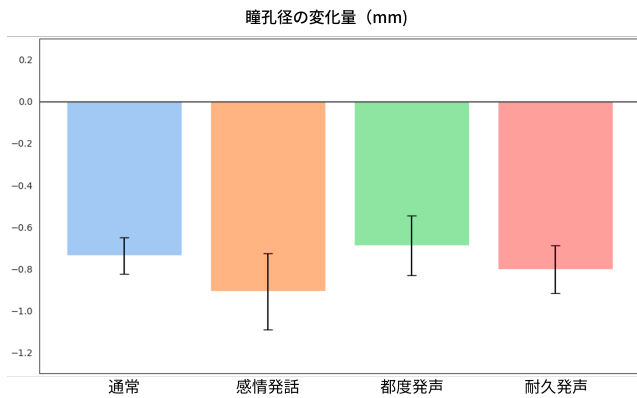


図 3 瞳孔径の変化量

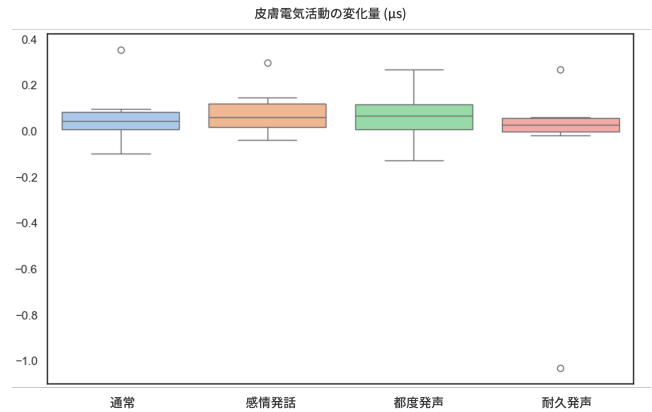


図 4 皮膚電気活動 (EDA) の変化量

いなどの意見を得た。都度発声を選んだ人からは驚きや急な状況変化に対する焦りなどのとっさに出る大きな感情変化や反射で感情を表現しやすい点においてやりやすかったという意見を得た。ゲームへの集中は4人が耐久発声、4人が都度発声が最もタスクに集中できたと回答した。耐久発声を選んだ人からは発声し続けることによって何か感情が起こっても同じように集中してプレイし続けられるという意見を得た。都度発声を選んだ人からは発声をしないため発声を意識することなくタスクに集中できたとの意見を得た。

4. 議論

4.1 あアラウド法における発声方法

実験の結果、あアラウド法は実際に感情を言葉にするよりも小さな負荷で自身の感情を表現できることが明らかになった。また、あアラウド法の発声方法ごとに主観的な作業負荷に大きな差がないことが示唆された。加えて、タスク中の感情表現がしやすいと答えた人が最も多かった。さらに、高山らの実験 [4] では、耐久発声が都度発声に比べて、感情の生起を発見しやすく、感情の分析しやすいため、分析者視点では耐久発声が最も優れているとした。これらのことから、実験参加者は感情表現がしやすく、分析者は感情の分析がしやすいため、あアラウド法における発声方法は耐久発声が最も適している可能性がある。

4.2 「あ」以外の感嘆詞との組み合わせ

発声する言葉を変化させることで、より詳細な感情評価ができる可能性がある。実験後のインタビューにおいて、都度発声の際に「あ」以外の「お」や「え」、ハミングといった発声が出てしまったとの意見があり、実際にそのように発声してしまった実験参加者がいた。これまでのあアラウド法では、発声を感嘆詞である「あ」のみに制限することで、言語化の負担を減らし直接的な感情の観測する。これは川島らが驚きや感嘆などの快感と落胆や否定などの不快感情の両方を表現できること、「お」や「ん (ハミング)」

と比較して発声負担が変わらないことから、発声を「あ」のみに制限した。一方で、「お」や「ん (ハミング)」などと組み合わせることで、より快度の感情分析がおこないやすくなる可能性がある。川島らの実験 [31] で、「お」はポジティブな感情が表現しやすく、「ん (ハミング)」は悩みなどのネガティブな感情を表現しやすかったとのアンケート結果を得た。このことから、発声する感情ごとに表現しやすい感情が異なる可能性がある。これまであアラウド法では、条件を統一するために「あ」のみを発声させたが、他の発声方法を組み合わせることで快度をより、分析しやすくなる可能性がある。一方で、「お」や「ん (ハミング)」を始めとした感嘆詞が分析に与える影響については分かっていない。そのため、今後は発声する感嘆詞が、感情表現や感情観測に与える影響を調べ、体験者と分析者の双方に適した方法を探索する。

4.3 他のゲームやコンテンツへの利用可能性と展望

あアラウド法は、音声とゲーム画面や体験中の画面を取得して分析を行うため、身体計測器や特別な環境を必要としない。そのため、ビデオゲームだけでなく、インタラクティブデバイスや VR コンテンツのような、さまざまなコンテンツにおいて感情評価ができる可能性がある。しかし、あアラウド法を利用するにはコンテンツによって向き不向きがある可能性がある。今回の実験では体験者の視点では発声方法ごとの違いはなかったものの、高山らの実験では体験者視点では都度発声が最も適しているとしている。これは評価の仕方の違い以外にも、タスクとして行ったビデオゲームが影響した可能性がある。他にも、体験者が発声し続けるため、長時間のコンテンツを評価には向かない可能性がある。よって、あアラウド法をエンタテインメントシステムの評価に用いるには、エンタテインメントシステムごとの特徴をふまえ、コンテンツ毎の具体的な評価基準やさらなる実験手順の整理が必要である。今後は、観測手法と分析手法の両面を改善し、実験手順を明確化する。さらに、さまざまなエンタテインメントシステムを用いたケー

ススタディから、本手法の利用に適したコンテンツの特徴や利用場面について調査を進める。

5. 終わりに

本論文では、ゲーム体験中の発声方法の違いが、体験者にどのような負荷を与えるかを調査した。実験ではビデオゲームをプレイしながら体験者が感じた感情を発話やアアラウド法を用いて表現させ、主観的な認知負荷と生体反応を調べた。実験の結果、アアラウド法は発話によって感情を表現することに比べて、認知負荷が低いことが示された。実験の結果から、アアラウド法の最適な発声や今後の展望について考察した。

参考文献

- [1] Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W. and Taylor, J. G.: Emotion recognition in human-computer interaction, *IEEE Signal processing magazine*, Vol. 18, No. 1, pp. 32–80 (2001).
- [2] Hazlett, R.: Measurement of User Frustration: A Biologic Approach, *CHI '03 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '03, New York, NY, USA, Association for Computing Machinery, p. 734–735 (online), DOI: 10.1145/765891.765958 (2003).
- [3] 川島拓也, 渡邊恵太: あアラウド法: 体験中のリアルタイムな感情の観測手法とその基礎検討, *情報処理学会論文誌*, Vol. 64, No. 11, pp. 1483–1492 (2023).
- [4] 高山英里, 能丸天志, 安中勇貴, 山岸丈留, 渡邊恵太: 「あアラウド法」における感情観測手法の改善と分析手法の検討, *エンタテインメントコンピューティングシンポジウム 2023 論文集*, Vol. 2023, pp. 52–61 (2023).
- [5] Bargas-Avila, J. A. and Hornbæk, K.: Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience, *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 2689–2698 (2011).
- [6] Hassenzahl, M.: The thing and I: understanding the relationship between user and product, *Funology 2: from usability to enjoyment*, pp. 301–313 (2018).
- [7] Mandryk, R. L., Inkpen, K. M. and Calvert, T. W.: Using psychophysiological techniques to measure user experience with entertainment technologies, *Behaviour & Information Technology*, Vol. 25, No. 2, pp. 141–158 (online), DOI: 10.1080/01449290500331156 (2006).
- [8] Wiebe, E. N., Lamb, A., Hardy, M. and Sharek, D.: Measuring engagement in video game-based environments: Investigation of the User Engagement Scale, *Computers in Human Behavior*, Vol. 32, pp. 123–132 (2014).
- [9] Hafeez, T., Umar Saeed, S. M., Arsalan, A., Anwar, S. M., Ashraf, M. U. and Alsubhi, K.: EEG in game user analysis: A framework for expertise classification during gameplay, *Plos one*, Vol. 16, No. 6, p. e0246913 (2021).
- [10] Nacke, L. E., Grimshaw, M. N. and Lindley, C. A.: More than a feeling: Measurement of sonic user experience and psychophysiology in a first-person shooter game, *Interacting with computers*, Vol. 22, No. 5, pp. 336–343 (2010).
- [11] Kallinen, K.: The effects of background music on using a pocket computer in a cafeteria: Immersion, emotional responses, and social richness of medium, *CHI'04 Extended Abstracts on Human Factors in Computing Systems*, pp. 1227–1230 (2004).
- [12] Reason, M. and Reynolds, D.: Kinesthesia, Empathy, and Related Pleasures: An Inquiry into Audience Experiences of Watching Dance, *Dance research journal*, Vol. 42, No. 2, pp. 49–75 (2010).
- [13] Bernhaupt, R.: Evaluating User Experience in Games Concepts and Methods, *London, UK* (2010).
- [14] Ekman, P. and Friesen, W. V.: Constants across cultures in the face and emotion., *Journal of personality and social psychology*, Vol. 17, No. 2, p. 124 (1971).
- [15] Banse, R. and Scherer, K. R.: Acoustic profiles in vocal emotion expression., *Journal of personality and social psychology*, Vol. 70, No. 3, p. 614 (1996).
- [16] Barsoum, E., Zhang, C., Ferrer, C. C. and Zhang, Z.: Training deep networks for facial expression recognition with crowd-sourced label distribution, *Proceedings of the 18th ACM international conference on multimodal interaction*, pp. 279–283 (2016).
- [17] Tzirakis, P., Trigeorgis, G., Nicolaou, M. A., Schuller, B. W. and Zafeiriou, S.: End-to-end multimodal emotion recognition using deep neural networks, *IEEE Journal of selected topics in signal processing*, Vol. 11, No. 8, pp. 1301–1309 (2017).
- [18] Damasio, A. R.: *The feeling of what happens: Body and emotion in the making of consciousness*, Houghton Mifflin Harcourt (1999).
- [19] Kreibitz, S. D.: Autonomic nervous system activity in emotion: A review, *Biological psychology*, Vol. 84, No. 3, pp. 394–421 (2010).
- [20] Gwin, J. T., Gramann, K., Makeig, S. and Ferris, D. P.: Removal of movement artifact from high-density EEG recorded during walking and running, *Journal of neurophysiology*, Vol. 103, No. 6, pp. 3526–3534 (2010).
- [21] for Psychophysiological Research Ad Hoc Committee on Electrodermal Measures, S., Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., Roth, W. T., Dawson, M. E. and Filion, D. L.: Publication recommendations for electrodermal measurements, *Psychophysiology*, Vol. 49, No. 8, pp. 1017–1034 (2012).
- [22] Watson, D., Clark, L. A. and Tellegen, A.: Development and validation of brief measures of positive and negative affect: the PANAS scales., *Journal of personality and social psychology*, Vol. 54, No. 6, p. 1063 (1988).
- [23] Bradley, M. M. and Lang, P. J.: Measuring emotion: the self-assessment manikin and the semantic differential, *Journal of behavior therapy and experimental psychiatry*, Vol. 25, No. 1, pp. 49–59 (1994).
- [24] Izard, C. E. and Read, P. B.: *Measuring emotions in infants and children: based on seminars sponsored by the Committee on Social and Affective Development During Childhood of the Social Science Research Council*, Vol. 1, Cambridge University Press (1982).
- [25] O'Brien, H. L. and Toms, E. G.: The development and evaluation of a survey to measure user engagement, *Journal of the American Society for Information Science and Technology*, Vol. 61, No. 1, pp. 50–69 (2010).

- [26] Ericsson, K. A. and Simon, H. A.: Protocol analysis: Verbal reports as data (Revised Edition) (1993).
- [27] Kato, T.: What “question-asking protocols” can say about the user interface, *International Journal of Man-Machine Studies*, Vol. 25, No. 6, pp. 659–673 (1986).
- [28] 長谷川凌, 川島朋也, 篠原一光: 思考発話法における発話量の個人差とワーキングメモリ容量の関係, *認知心理学研究*, Vol. 21, No. 2, pp. 47–58 (オンライン), DOI: 10.5265/jcogpsy.21.47 (2024).
- [29] Zhang, T., El Ali, A., Wang, C., Hanjalic, A. and Cesar, P.: RCEA: Real-time, Continuous Emotion Annotation for Collecting Precise Mobile Video Ground Truth Labels, *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, New York, NY, USA, Association for Computing Machinery, p. 1–15 (online), DOI: 10.1145/3313831.3376808 (2020).
- [30] Iida, K., Ina, Y., Hayashi, D., Yanase, Y. and Watanabe, K.: Open Video Game Library: Developing a Video Game Database for Use in Research and Experimentation, *Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology*, VRST '23, New York, NY, USA, Association for Computing Machinery, (online), DOI: 10.1145/3611659.3617194 (2023).
- [31] 川島拓也, 渡邊恵太: あアラウド法: 体験中のリアルタイムな感情の観測手法とその基礎検討, *情報処理学会論文誌*, Vol. 64, No. 11, pp. 1483–1492 (2023).