

あアラウド法における表情の有無が感情評価に与える影響の調査

藤田 のどか^{1,a)} 大見 亮介^{1,b)} 山岸 丈留² 安中 勇貴² 能丸 天志² 渡邊 恵太^{1,c)}

概要：タスクをしながら頭に思い浮かんだことを語る思考発話法では、ユーザの心理プロセスを獲得できる。しかし、言語化の負担やユーザによる語彙の差、データの収集とその評価が難しいなどの問題がある。これに対して、川島らは感嘆詞「あ」の発声を用いて体験中の感情をリアルタイムに観測する手法、「あアラウド法」を提案した。しかし、あアラウド法は体験中の声とゲーム画面のみから感情を推定しており、分析者間での一致度には改善の余地がある。そこで本研究では、あアラウド法の分析対象に体験中の表情を加えることで、覚醒度および快度の一致度が向上するかを検証する。結果、覚醒度においては表情を加えたことによる大きな差はなく、快度においては表情を加えた場合に一致度が向上した。議論では、実験結果から明らかになった課題を踏まえ、分析対象の選定について言及し、あアラウド法のさらなる精度向上に向けた今後の研究の方向性を示した。

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、感情変容の観測は重要な役割を果たす [1][2]。水口ら [3] は、映画やゲーム、スポーツなどの日常的なエンタテインメント事例を通じて、エンタテインメントが感情の変化と深く関わることを示した。これまで、感情観測のためにリッカートスケールやイラストを用いたアンケートが複数開発されている [4][5][6]。また、アンケートに加え、実験後インタビューによる感情報告も多くの研究で実施される [7]。しかし、これらの方法には忘却や捏造の可能性がある。また、事後評価のためシステムの全体的な評価にとどまり、細部の改善に活用することは難しい [8]。

こうした課題の中で、表情を用いた感情評価が注目されている。多くの研究が、表情分析が感情の推定において重要な情報を提供することを示している [9]。例えば、Murugappan ら [10] は顔の幾何学的特徴を抽出し、感情表現を分類するための機械学習アルゴリズムの性能を評価した。その結果、幾何学的特徴は感情表現の分類において有効であることを明らかにした。一方で、データセットの多様性や品質が分類精度に影響を与えることを示唆した。ま

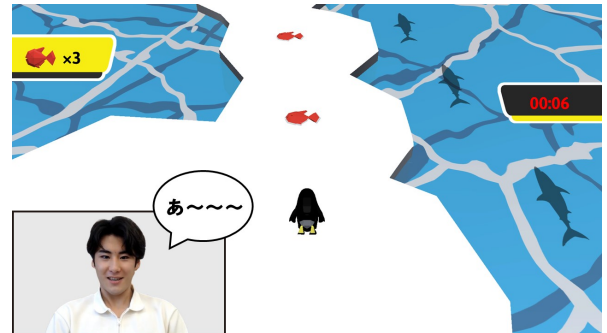


図 1 あアラウド法：タスク中の感情を「あ」の発声を用いて表現し、その抑揚や大きさから感情を推測する手法。

た、表情分析にはリアルタイムでの計測やノイズの影響、文化や個人による差が大きいなどの指摘がある [11]。

以上の理由から、感情観測の手法には依然として課題が多く、エンタテインメントシステムの精緻な評価や改善に向けた新たなアプローチが求められている。

そこで川島ら [12] は、システム体験中の感情をリアルタイムに得る手法、「あアラウド法」を提案した (図 1)。この手法では、実験参加者が感嘆詞「あ」を用いて感情を表現し、その声の抑揚や大きさを分析することで感情を観測する。川島らは、あアラウド法が従来の思考発話法に比べて参加者の負担が少なく、多様な感情や思考を「あ」で表現できることを明らかにした。山岸ら [13] は、あアラウド法により取得したデータを用いて、Russell の円環モデル [14] を基に開発された Affect Grid [15] の枠組みを活用し、時系列的な感情変容を分析した。実験参加者の覚醒度と快度を

¹ 明治大学 総合数理学部 先端メディアサイエンス学科

² 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

a) ev220505@meiji.ac.jp

b) ev220600@meiji.ac.jp

c) keita_w@meiji.ac.jp

1 秒ごとに評価し、評価者間の一致度を確認したところ、一致度は約 67 % でカッパ係数の値は 0.95 を超える結果となった。一方で、山岸らの研究では声とゲームプレイ画面を分析対象としていたが、表情は含まれていなかった。表情を用いた感情分析が多くの研究で実施されることから、表情を分析対象に加えることで、あアラウド法の精度が向上する可能性がある。

そこで本研究では、体験中の表情を分析対象に加えて、山岸らと同様の実験を行う。本研究の目的は、表情を用いることで、あアラウド法における評価者間の一致度が向上するか検証することである。実験では級内相関係数 (ICC) を用いて、覚醒度と快度の一致度を分析した。実験結果から明らかになった課題を踏まえ、分析対象の選定について言及し、あアラウド法のさらなる精度向上に向けた今後の研究の方向性を示した。本研究の結果が、感情推定のさらなる発展と、感情分析分野全体の技術向上に寄与することを目指す。

2. 関連研究

2.1 HCI 分野における感情評価手法と課題

感情は体験評価において重要な要素の一つであり [8][16][17][18]、モバイルアプリやウェブサイトなど、様々なコンテンツやシステムで感情を指標にしている。既存の感情指標は大きく分けて、主観指標、生理指標、行動指標である [19]。Javier ら [8] によると、HCI 分野では半数以上の研究で主観指標を用いる。主観指標は、アンケートやインタビュー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。Norman ら [20] は、体験評価はインタラクション中の経験自体でなく、経験に関する記憶が重要であると述べた。一方で、事後報告には忘却や捏造のリスクがあるうえ、得られるデータは質問項目や実験参加者の言語化能力に依存する。また、アンケートやインタビューでは時系列な感情変容を得られず、回顧評価とシステム体験との相互作用や、システム体験の詳細の確認が難しい [8]。そこで、生理指標や行動指標をアンケートやインタビューと併用し、感情観測の信頼性、妥当性を高める手法を検討している [21][22]。一方で、アンケートへの回答と生体反応には有意な不一致が存在するという指摘があり [23]、ロバストな評価手法を確立できていない。

2.2 表情を用いた感情評価手法

表情は、様々な行動指標の中でも感情との関連が特に調査されている。例えば、機械学習により顔の動きや特徴点を解析することで感情の自動検出を行う研究や [9]、特定の筋肉の動きを捉え、「喜び」「悲しみ」などの感情カテゴリーに分類する研究がある [10]。一方で、表情による感情表現は必ずしも普遍的ではなく、文化差や個人差が大きい

という指摘がある [11]。例えば、同じ「喜び」の感情でも、欧米文化圏では笑顔が顕著に表れることが多いが、アジア文化圏では控えめな表情であることが多い。また、個人によっても表情の動き方や強さが異なるため、単純な表情解析だけでは感情を評価することは難しい。これらの問題を克服するために、表情認識と生体データを組み合わせたマルチモーダルな感情評価手法の提案がある [24]。このアプローチにより、表情の個人差や文化差を補完する、より高精度な感情検出の実現が期待できる。

2.3 あアラウド法

あアラウド法は、実験参加者がシステム体験中の感情を「あ」のみで表現し、その抑揚や大きさから感情をリアルタイムに観測する手法である。先行研究では、「あ」は日本語のなかでも日常生活によく用いられる感嘆詞の一つであり、多様な心理プロセスを表現できることを示唆した [25]。川島ら [26] は、「あ」の音声的特徴を通じて 7 種類の思考や感情を表現できることを確認した。高山ら [27] は、ゲーム中の実験参加者の「あ」の発声とゲーム画面の録画から、第三者が感情生起のタイミングを判別し喜怒哀楽に分類できることを確認した。これらの調査は、あアラウド法によって実験参加者の感情変容を観測できる可能性を示唆した。また、安中ら [28] は、身体負荷は無言、自由発話、あアラウド法のいずれの条件にも有意差がないこと、認知負荷は自由発話の場合のみ高いことを確認した。この結果は、「あ」のみで感情を表現する行為そのものが、体験中や事後の評価に与える影響が少ないことを示唆している。

3. 実験

本研究では、あアラウド法を行いながらビデオゲームをプレイした動画を使用し、「あ」の発声から感情を分析する実験を行った。分析対象には、「あ」の発声とゲーム画面に加えて、ゲームプレイ中の表情を含めた。実験参加者は大学生計 6 名で、実験はすべてオンラインで実施した。

3.1 実験参加者

実験参加者は平均年齢 20 歳の大学生計 6 名（男性：5 名、女性：1 名、年齢：19～21 歳）である。プライバシーを厳重に保護するため、取得したすべての個人情報は匿名化して管理した。データは研究者間でのみ利用し、外部の第三者とは一切共有しなかった。

3.2 実験条件 感情評価シート

感情評価には Russell の円環モデル [14] を応用して開発された Affect Grid [15] の枠組みを活用し実験した。Affect Grid は、ラッセルの円環モデルを簡略化した評価ツールで

ある。覚醒度と快度の2軸に基づいて、9×9のグリッド上に点をプロットし、感情状態を反映する。数値の記入のしやすさと分析のしやすさの観点から、分析シートはExcel形式で作成した(図2)。行にはそれぞれ覚醒度と快度、列には残り秒数を示し、各マス目に1秒ごとの覚醒度と快度を1~9の9段階で記入した。覚醒度は1のときに眠気がある状態、5が中立、9が最も覚醒している状態とした。快度は1のときに不快な状態、5が中立、9が最も快適な状態とした。また、息継ぎにより「あ」の発声がない秒数は著者が事前に確認し、その秒数の分析は行わなかった。

評価対象の動画

実験では、感情評価の対象にゲームプレイ中の動画を使用した。共同研究者(男性1名)がプレイ動画と基準音声を収録した。あアラウド法を用いてゲームをプレイした際の音声とゲーム画面、共同研究者の表情を複数の組み合わせで収録し、計4つの動画を作成した。基準音声とは、覚醒度と快度が共に5の場合の「あ」の発声である。収録した基準音声は4つの動画すべての冒頭に合成し、実験参加者が感情を評価する際の参考になるようにした。ゲームは飯田らのOpen Video Game Library[29]の中のビデオゲーム「Sliding Penguin」[30][31]を利用した。

実験で使用した動画の撮影は以下の手順で行った。

- (1) 撮影の概要説明
- (2) 発声方法の説明
- (3) 使用するゲームの説明
- (4) ゲーム操作と感情表現の練習
- (5) 基準音声の収録
- (6) ゲームプレイ動画の撮影

撮影時の環境音が実験参加者の評価に影響を与えないよう、撮影環境は静かな場所を選定した。「あ」の発声方法は、山岸ら[13]と同様の手法である耐久発声を採用した。ゲームの制限時間は30秒に設定した。ゲーム性を高め感情を出やすくするためにゲームステージ上に魚を配置し、プレイ中はなるべく多くの魚を取るように指示した。プレイ動画は以下に記載の通り全部で4パターンを収録し、パターンによって動画に含まれる項目を変更した。

- 動画1. 「あ」の発声+ゲーム画面
- 動画2. 「あ」の発声+ゲーム画面+表情
- 動画3. 「あ」の発声+表情
- 動画4. 「あ」の発声

3.3 実験環境

実験はすべてオンラインで実施した。静かな場所で動画の音声がはっきりと聞こえ、乱れや遅延なく動画を再生できるスペックのPCとネット環境の下で実施した。実験参加者はイヤホンまたはヘッドホンを装着して周囲の騒音を遮断し、動画の音声に集中可能な状態にした。また、

		この部分はプロットしなくて大丈夫です！								
	秒数	0:29	0:28	0:27	0:26	0:25	0:24	0:23	0:22	0:21
覚醒度(1:眠気－9:覚醒)										
快度(1:不快－9:快)										
判断に迷ったなど何かあればメモ										

図2 感情分析シート。行にはそれぞれ覚醒度と快度、列には残り秒数を示し、分析しなくてよい箇所は著者が事前に赤く塗りつぶした。

Bluetoothによる音声の遅延を防ぐため、イヤホンまたはヘッドホンは有線のものを使用するように指示した。

3.4 実験手順

実験参加者の実験の流れを以下に示す。

- (1) 実験同意書を記入する
- (2) 実験概要についての説明資料を読む
- (3) 覚醒度・快度の説明、評価方法と注意点についての説明資料を読む
- (4) 実験環境を整える
- (5) 実験タスクを行う
- (6) 事後アンケートに答える

実験参加者は送付された動画とシートを用いて、動画中の「あ」に対応する感情を分析した。約30秒間の動画の1秒ごとの覚醒度・快度をそれぞれ1-9の9段階で評価し、その値を評価シートに記録した。1秒間に複数の感情が表出した場合は、一番長い時間出ていたと感じる感情を評価した。動画は巻き戻しが可能であり、実験参加者は納得がいくまで繰り返し視聴した。動画に表示されるゲーム画面や共同研究者の表情は、評価をする際の参考にしてよいものとした。実験参加者はこれらの手順を、実験者が用意した4つの動画それぞれについて、動画1から4の順番で繰り返し実施した。

また事後アンケートでは、年齢・性別、健康状態、あアラウド法の発声・分析経験、Sliding Penguinのプレイ経験、ゲーム習慣について回答した。また、分析が難しいと感じた箇所があればその理由や、実験全体を通しての感想についても回答した。

4. 結果と考察

4.1 覚醒度と快度の評価の一致度

本研究では、複数の実験参加者間の一致度を評価するため、級内相関係数(ICC)を採用した。中でも、複数の実験参加者が複数の対象(動画)を一回ずつ評価する今回の実験デザインに適したICC2k(絶対一致)を使用した。

表1に各動画における覚醒度と快度のICC値、p値、CI95%の範囲を示す。覚醒度のICC値および95%信頼区間は、「発声+ゲーム画面+表情」、「発声+ゲーム画面」、「発声」、「発声+表情」の順番で高い結果となった。しか

し、ICC は 4 つの動画すべてにおいて 0.87 以上を示し、95%信頼区間の下限もすべての動画において 0.79 以上であったことから、覚醒度に関してはどの動画においても高い一致度を得られていることを示した。

一方、快度の ICC 値および 95%信頼区間は、「発声＋表情」、「発声」、「発声＋ゲーム画面＋表情」、「発声＋ゲーム画面」の順番で高い結果となった。特に、「発声＋ゲーム画面」においては ICC 値が 0.159580、95%信頼区間の下限が-0.25 と 4 つの動画の中で唯一マイナスの値になった。この結果は、評価対象にゲーム画面を含んだ場合に一致度が低下する可能性があることを示唆した。

4.2 事後アンケート結果

実験後のアンケートでは、以下の回答を得た。

- (1) 「発声＋ゲーム画面＋表情」のように情報量が増えると、どの情報から評価をしたらよいのか迷った。
- (2) 「発声＋ゲーム画面」を視聴したときは楽しそうにプレイしてるように感じたが、表情が含まれる「発声＋ゲーム画面＋表情」と「発声＋表情」を視聴してからは苛立ちを感じながらプレイしているように感じたため、表情の影響は大きいと感じた。
- (3) 「発声＋ゲーム画面＋表情」と「発声＋表情」で表情を見た後に「発声」を評価したため、記憶の中の表情が「発声」の評価に影響してしまった。動画を視聴する順番によって分析結果が大きく変わりそうと思った。
- (4) 評価する秒数の直前の声や表情、プレイ画面を考慮するか迷った。今回は評価する秒数にのみ注目するようにした。
- (5) 今までに複数回あアラウド法の分析経験があるが、共同研究者の感情表現が上手だったため、今までの中で最も評価が簡単だった。

アンケート結果より、覚醒度においてすべての動画で高い一致度を得たのは、共同研究者の感情表現が上手であったからだという可能性もある。また、快度において「発声＋表情」の一致度が最も高い結果となったことから、表情を含むと特定の感情を感じやすい可能性がある。

5. 議論

5.1 感情評価における表情の影響

快度の評価において、「発声＋表情」が最も高い一致度を示したことから、表情の情報が感情評価に寄与する可能性がある。例えば、ゲーム終了直前において、「発声＋ゲーム画面」では快度を 8 や 9 (平均値 6.3) と評価する実験参加者が多かった。一方、「発声＋表情」では 1 や 2 (平均値 2.5) が多く、「発声＋ゲーム画面＋表情」では評価が大きく分かれ、平均値は 4.2 となった。また、各動画のゲーム終了直前の発声を著者が確認したところ、いずれの動画に

表 1 動画ごとの覚醒度・快度の ICC 値、p 値及び CI95%の範囲

*は $p < 0.05$ の値であり、分析結果の一致度が偶然によるものでなく統計的に有意な数値であることを示す。

動画	尺度	ICC	p 値	CI95%
動画 1	覚醒度	0.906302	1.417998e-26*	[0.82, 0.96]
	快度	0.159580	0.204687	[-0.25, 0.53]
動画 2	覚醒度	0.922456	6.257649e-26*	[0.87, 0.96]
	快度	0.596881	0.000299*	[0.32, 0.79]
動画 3	覚醒度	0.874806	1.839429e-18*	[0.79, 0.93]
	快度	0.794046	4.390799e-12*	[0.65, 0.89]
動画 4	覚醒度	0.891517	6.846391e-21*	[0.81, 0.95]
	快度	0.732108	8.186624e-08*	[0.54, 0.86]

においても「あ」の発声に大きな違いはなかった。この結果から、実験参加者は表情を参考に快度を低く評価した可能性がある。

さらに、共同研究者は「あ」を発声しているため、ずっと口を開けている表情となるが、その中でも口の開き方に違いがあった。例えば、多くの実験参加者が快度を不快側に評価した秒数では、口の端をゆがめるような表情をしていた。このことから、発声中の微妙な表情の変化が評価に影響を与えている可能性がある。また、あアラウド法において表情を含む分析手法が感情推定の精度向上に寄与する可能性を示した。

事後アンケートでは、「発声＋ゲーム画面」を視聴したときは楽しそうに感じたが、表情が含まれる「発声＋ゲーム画面＋表情」と「発声＋表情」を視聴してからは苛立っているように感じたという意見があった。この結果は、声やゲーム画面では捉えられない感情を表情から読み取れる可能性を示唆した。ただし、別の回答では、表情の追加による情報量の増加が分析者の負担を増大させるという指摘もあった。

5.2 あアラウド法における適切な評価対象

覚醒度の一致度に関して、4 つの動画すべてにおいて ICC 値が 0.87 以上、95%信頼区間の下限も 0.79 と高い結果になったことから、覚醒度に関しては、動画に含まれる要素に関わらず高い一致度を得ることを示唆した。このことから、ゲーム画面や表情を除いた「あ」の発声のみでも評価ができる可能性がある。あアラウド法は今まで、「あ」の発声とプレイ中のゲーム画面を用いて評価を行っていたが、声のみで評価をすることにより負担が軽減する可能性がある。

一方、快度に関しては、「発声＋ゲーム画面」が最も低い一致度、次いで「発声＋ゲーム画面＋表情」が低い一致度となった。この結果から、評価対象にゲーム画面を含んだ場合に一致度が低下する可能性がある。ゲーム画面はプレイ状況が明確に理解できる一方で、快度の評価には不適合な可能性がある。ゲーム画面をどのくらい考慮したかに

よって分析者間での評価が大きく異なり、一致度が下がる要因になった可能性がある。

5.3 分析の順番や慣れの影響

事後アンケートから、動画を視聴する順番が評価に影響を及ぼすことを示唆した。実験参加者のうち1名は、表情を含む「発声+ゲーム画面+表情」と「発声+表情」の視聴後に「発声」を視聴した。そのため、表情が含まれないパターンであるにもかかわらず、直前に見ていた表情に影響されてしまったと述べた。今回、「あ」の発声のみで構成されている動画が覚醒度、快度のどちらにおいてもある程度の一貫性を得ていたのは、視聴順序が最後だったことに起因する可能性がある。

また、今までに複数回あアラウド法の分析者として評価を行ったことのある実験参加者は、共同研究者の感情表現が上手だったため、今までの中で最も評価が簡単だったと述べた。覚醒度の評価がどの動画でも高い結果になったことや、快度においても3つの動画でICC値が0.59を上回る高い一貫性となった理由は、今回ゲームをプレイした共同研究者が、感情表現に優れているからだという可能性がある。

6. おわりに

本論文では、あアラウド法の Affect Grid による感情推定において、分析対象に表情を加えることによる効果を調査した。

実験結果から、表情を加えることで感情推定の一貫性が向上する可能性を示唆した。一方で、情報量の増加により分析者の負担が増大する課題も明らかとなった。覚醒度の評価においては、動画間で一貫性に大きな差はなく、声のみでも評価ができる可能性がある。また、快度の評価においては、ゲーム画面を含む場合に一貫度が低下する傾向があったことから、各要素の役割や適切な組み合わせについてさらなる検討が必要である。

今後は、分析者の負担を軽減しつつ感情推定の精度を向上させるために、分析対象の最適な組み合わせや提示方法を体系的に検証する必要がある。本研究で得られた知見をもとに、異なる状況やタスクにおけるあアラウド法の適用可能性を広げるための応用研究を進めていく。

参考文献

- [1] Roddy Cowie, Ellen Douglas-Cowie, Nicolas Tsapatsoulis, George Votsis, Stefanos Kollias, Winfried Fellenz, and John G Taylor. Emotion recognition in human-computer interaction. *IEEE Signal processing magazine*, Vol. 18, No. 1, pp. 32–80, 2001.
- [2] Richard Hazlett. Measurement of user frustration:

- A biologic approach. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '03, p. 734–735, New York, NY, USA, 2003. Association for Computing Machinery.
- [3] 水口充, 片寄晴弘. エンタテインメントコンピューティング研究における評価問題の解決に向けての施策の実践. 情報処理学会 インタラクシオン 2019, pp. 141–150, 2019.
- [4] David Watson, Lee Anna Clark, and Auke Tellen. 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.
- [5] Albert Mehrabian. Framework for a comprehensive description and measurement of emotional states. *Genetic, social, and general psychology monographs*, Vol. 121 3, pp. 339–61, 1995.
- [6] Carroll Ellis Izard and Peter B Read. *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.
- [7] Ingrid Pettersson, Florian Lachner, Anna-Katharina Frison, Andreas Riener, and Andreas Butz. A bermuda triangle? a review of method application and triangulation in user experience evaluation. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, pp. 1–16, 2018.
- [8] Javier A Bargas-Avila and Kasper Hornbæk. Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 2689–2698, 2011.
- [9] S. asif Hussain and Ahlam Al Balushi. A real time face emotion classification and recognition using deep learning model. *Journal of Physics: Conference Series*, Vol. 1432, p. 012087, 01 2020.
- [10] Murugappan M. and Mutawa A. Facial geometric feature extraction based emotional expression classification using machine learning algorithms. *PLOS ONE*, Vol. 16, No. 2, pp. 1–20, 02 2021.
- [11] JA Russell. Is there universal recognition of emotion from facial expression? a review of the cross-cultural studies. *Psychological bulletin*, Vol. 115, No. 1, p. 102–141, January 1994.
- [12] 川島拓也, 渡邊恵太ほか. 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.
- [13] Takeru Yamagishi, Yuki Yasunaka, Eri Takayama, Takashi Nomaru, Takanori Komatsu, and Keita Watanabe. Ah-aloud method to comprehend time-series emotion observation during gameplay: An initial investigation with japanese speakers. *Proceedings of the ACM on Human-Computer Interaction*, Vol. 8, No. CHI PLAY, pp. 1–23, 2024.
- [14] James A Russell. A circumplex model of affect. *Journal of personality and social psychology*, Vol. 39, No. 6, p. 1161, 1980.
- [15] James A Russell, Anna Weiss, and Gerald A Mendelsohn. Affect grid: a single-item scale of pleasure and arousal. *Journal of personality and social psychology*, Vol. 57, No. 3, p. 493, 1989.
- [16] Manfred Thüring and Sascha Mahlke. Usability, aesthetics and emotions in human–technology interaction. *International journal of psychology*, Vol. 42, No. 4, pp. 253–264, 2007.
- [17] Don Norman. *Emotional design: Why we love (or hate) everyday things*. Basic books, 2007.
- [18] Pertti Saariluoma and Jussi Jokinen. Emotional dimensions of user experience: A user psychological analysis. *International Journal of Human-Computer Interaction*, Vol. 30, , 04 2014.
- [19] Seth Kaplan, Reeshad Dalal, and Joseph Luchman. *Measurement of emotions.*, pp. 61–75. 01 2013.
- [20] Donald A. Norman. The way i see itmemory is more important than actuality. *Interactions*, Vol. 16, No. 2, p. 24–26, mar 2009.
- [21] Irene Lopatovska and Ioannis Arapakis. Theories, methods and current research on emotions in library and information science, information retrieval and human-computer interaction. *Inf. Process. Manage.*, Vol. 47, pp. 575–592, 07 2011.
- [22] Anders Bruun, Effie Lai-Chong Law, Matthias Heintz, and Poul Svante Eriksen. Asserting real-time emotions through cued-recall: is it valid? In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*, pp. 1–10, 2016.
- [23] Anders Bruun and Simon Ahm. Mind the gap! comparing retrospective and concurrent ratings of emotion in user experience evaluation. Vol. 9296, 09 2015.
- [24] 池田悠平, 岡田佳子, 堀江亮太, 菅谷みどり. 表情と生体情報を用いた感情の推測方法の検討. 分散協調とモバイルシンポジウム, Vol. 2016, pp. 149–161, 2016.
- [25] Takeshi Kohno. [emotives and intonation orchestrated in japanese: Affective relevance modality] nihongo ni okeru kandōshi to intonēshon no kōkyō : kanrensei modarithi no fuzei. *Otsuna Women's University annual report. Humanities and social sciences*, Vol. 51, pp. 226–207, 2019.
- [26] Takuya Kawashima and Keita Watanabe. [ah-aloud method: Proposal and validation of a method to evaluate psychological processes during an experience using the phonetic information of 'a'] a araudō hō: taiken chū no'shinri purosesu wo 'a' no onsei jōhō de hyōka suru shuhō no teian to kenshō. *Proceedings of Entertainment Computing Symposium 2022*, Vol. 2022, pp. 178–183, aug 2022.
- [27] Eri Takayama, Takashi Nomaru, Yuki Yasunaka, Takeru Yamagishi, and Keita Watanabe. [improvement of emotional observation method and analysis method in the "ah-aloud" method] a araudō hō ni okeru kannjou kansoku syuhō no kaizen to bunseki syuhō no kentō. In *Proceedings of Entertainment Computing Symposium 2023*, Vol. 2023, pp. 52–61, aug 2023.
- [28] 安中勇貴, 能丸天志, 山岸丈留, 渡邊恵太ほか. 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. エンタテインメントコンピューティングシンポジウム 2024 論文集, Vol. 2024, pp. 199–204, 2024.
- [29] Takuya Oka, Takuya Kawashima, Daichi Hayashi, and Keita Watanabe. [design and development of video games for ease of use and standardization of research] kenkyū riyō shi yasuku hyōjun sei wo mezashita bideo gēmu no sekkei to kaihatsu. *Proceedings of Entertainment Computing Symposium 2021*, Vol. 2021, pp. 181–186, aug 2021.
- [30] 岡拓也, 川島拓也, 林大智, 渡邊恵太ほか. 研究利用しやすく標準性を目指したビデオゲームの設計と開発. エンタテインメントコンピューティングシンポジウム論文集, Vol. 2021, pp. 181–186, 2021.
- [31] 岡拓也, 川島拓也, 築瀬洋平, 渡邊恵太ほか. Hci 研究における評価実験用ビデオゲームの要件探究とオープンビデオゲームライブラリを用いたケーススタディ. 研究報告ヒューマンコンピュータインタラクション (HCI), Vol. 2022, No. 12, pp. 1–8, 2022.