JAIST Repository
https://dspace.jaist.ac.jp/

Title	エージェント化プロトタイプとの対話を利用したユーザテスト手法の検討
Author(s)	吉松，駿平
Citation	
Issue Date	2023-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/18277
Rights	
Description	Supervisor: 高島 健太郎，先端科学技術研究科，修士(知識科学)

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

修士論文

エージェント化プロトタイプとの対話を利用したユーザテスト手法の検討

2110187　吉松　駿平

主指導教員　高島　健太郎

北陸先端科学技術大学院大学
先端科学技術研究科
（知識科学）

令和 5 年 2 月

Abstract

Examination of user testing methods using interaction with agent-based prototypes

Shunpei Yoshimatsu

In user-centered design methods such as design thinking, during the user testing phase, it is necessary to obtain the user experience when the prototype is used in order to improve the quality of the user experience. Therefore, in order to obtain the user's thoughts and emotions, Think Aloud method is used in the field of user testing. However, this method may not be able to obtain a sufficient amount of speech due to the high cognitive load. Additionally, developer intervention to support speech may cause the user to think too much, preventing intuitive speech and the risk of not obtaining honest reactions and impressions. In this study, we propose a method of adding physical characteristics to the prototype and making it an agent to encourage honest user speech and extract more usability improvement hints. We also conducted a comparison experiment with human developer intervention.

Through the examination of usability by analyzing the results of impression evaluation, speech collection, and qualitative analysis of speech data, subjective impressions of the agent-based prototype, such as "Ease of speaking" and "Type of relationship" varied among users. In the agent condition, all users' speech increased as a result of more verbal interaction than in the human condition where the developer sat next to them. As a result, it suggests that the respondents were encouraged to mention usability and were more likely to verbalize their discomfort with the functions and operation of the application. In addition, the increase in the number of responce was due to the ability to respond regardless of the user's mood,

which is considered to be an advantage unique to the Agent condition. Also, it was found that users in the Agent condition considered the responsibility for operation error to be on the interface side rather than their own operation. The content of speech also showed that users in the Agent condition expressed their frank impressions immediately, without making a logical storyline of argument or making any implication. However, users in the Agent condition tend to not speak in detail about their operation than the Human condition. Therefore, it is important to use not only the agent-based prototype, but also the developer's intervention when a more detailed explanation is needed.

目　次

第 1 章 はじめに 1

1.1 背景 1
1.2 問題提起 1
1.3 研究目的 3
1.4 論文の構成 3

第 2 章 関連研究 4

2.1 ユーザテストにおける TA 法の実施方法と介入の影響に関する研究 4
2.2 プロダクトやアプリケーションのエージェント化に関する研究 5

第 3 章 提案手法 6

3.1 提案手法の概要 6
3.2 セリフの選定・音声合成 6
3.3 身体的特徴の付与 8

第 4 章 実験 9

4.1 実験計画 9
4.2 測定内容 9
4.3 予備的実験 11
4.4 本実験 12

第 5 章 実験結果と考察 16

5.1 アンケート・インタビューに基づく分析 17
5.1.1 ユーザテスト中の「話しやすさ」の評価 18
5.1.2 被験者との「関係性」の評価 20
5.1.3 ユーザテストに関する項目の評価 24
5.1.4 アンケート・インタビューによる印象評価のまとめ 25

5.2 実験中の発話数・言及数に基づく分析 27
5.3 発話データに基づく分析 29
5.3.1 全体的な口調の違い 29
5.3.2 操作の説明の違い 30
5.3.3 責任の所在に関する表現の違い 32
5.3.4 ユーザビリティに関する言及・指摘時の違い 33
5.3.5 Human 条件特有の機能を褒める発言 34
5.3.6 各条件の発話内容のまとめ 36
第 6 章 結論 39
6.1 研究目的に対する回答 39
6.2 提案手法の有用性と各条件の使い分けの提案 39
6.3 本研究のまとめ 40
6.4 今後の展望 41
謝辞 42
付録 43
6.5 Human 条件のアンケート 43
6.6 Agent 条件のアンケート 46
参考文献 49

図　目　次

図 3-1　提案手法の画面イメージ.....8
図 4-1　実験に用いたペイントソフト　（上：muro，下：8bit paint）.....10
図 4-2　実験課題にしたイラスト.....11
図 4-3　予備実験中の実験室内　（被：被験者，実：実験者）.....12
図 4-4　本実験中の実験室内　（被：被験者，実：実験者，開：開発者）.....13
図 4-5　実験手順.....15
図 5-1　実験風景　（上：Human 条件，下：Agent 条件）.....16
図 5-2　録画された実験中の操作状況（Agent 条件）.....17
図 5-3　「考えたことや感じたことを素直に話せた」の被験者別評価.....18
図 5-4　「わがままだと思われそうな内容は思いついても話さなかった」の被験者別評価.....19
図 5-6　「開発者／キャラクターの存在を気にせずに話した」の被験者別評価..20
図 5-5　「開発者／キャラクターの機嫌が悪くなりそうな内容も話した」の被験者別評価.....20
図 5-7　「うまく操作できないときでもアプリ側の問題だと思えた」の被験者別評価.....21
図 5-8　「うまく操作できないとき自分のせいだと思った」の被験者別評価.....22
図 5-9　「開発者／キャラクターが喋ることで緊張した」の被験者別評価 22
図 5-10　「開発者／キャラクターは対等な関係だと感じた」の被験者別評価....23
図 5-11　「開発者／キャラクターは何を言っても構わない相手だと感じた」の 被験者別評価 24
図 5-12　「アプリの問題点を全て言うことができた」の被験者別評価.....25
図 5-13　1 分あたり発話数の被験者別変化.....27
図 5-14　1 分あたり言及数の被験者別変化.....28
図 5-15　言い放つような口調の場面の例（被験者 A, Agent 条件）.....29
図 5-16　開発者に答えを求める場面の例（被験者 D, Human 条件）.....30

図 5-17　自らの考えを開発者に確認する場面の例（被験者 B, Human 条件）...30
図 5-18　操作内容をそのまま発話している場面の例（被験者 C, Agent 条件）..31
図 5-19　丁寧な操作説明の例（被験者 B, Human 条件）.....32
図 5-20　「難しい」と発話した場面の例（被験者 B, Human 条件）.....32
図 5-21　機能がないことを指摘する場面（被験者 B, Agent 条件）.....33
図 5-23　違和感を感じたままに発話している場面の例（被験者 C, Agent 条件）
.....34
図 5-22　状況を整理してから発話している場面の例（被験者 C, Human 条件）
.....34
図 5-24　機能を褒める場面の例（被験者 C, Human 条件）.....34
図 5-25　開発者に指摘した直後にフォローしている場面の例（被験者 B, Human 条件）.....35
図 5-26　問題発生から発話までのプロセスの概要.....38

表　目　次

表 3-1　定型文の内容 7
表 4-1　予備的実験での 1 分あたりの発話数 12
表 4-2　実験計画（H：Human 条件，A：Agent 条件） 13
表 5-1　各被験者の 1 分あたりの発話数.....27
表 5-2　各被験者の 1 分あたりの言及数.....28
表 5-3　各条件の発話の特徴 37

第1章 はじめに

1.1 背景

近年，アプリケーションやプロダクト等の開発現場では，デザイン思考をはじめとしたユーザ中心の設計手法が多く用いられている．デザイン思考のアプローチは問題を解決するために，創造的で柔軟なアイデアを生み出すことを目的としており，次のようなステップからなるプロセスを通じて行われる．

まず，観察やインタビューから，問題点やニーズを明確にする．その後，問題を解決するための様々なアイデアを考え出す．生み出したアイデアを整理して最も有効なアイデアを選んだ後，具体的な形にプロトタイピングする．そして，作成したプロトタイプを使って，実際に問題を解決できるかどうかをテストし，そのテスト結果を踏まえてプロトタイプを改善する．

その中でも，プロトタイピングの成果物をユーザが試用し，その使用感や満足度を評価するテストはユーザテストと呼ばれる．ユーザに使ってもらうことで，プロトタイプが実際のユーザのニーズを満たしているかを確認することができる．また，一般的にユーザテストは繰り返し行われ，その結果を元にプロトタイプの改善点を特定し，修正を重ねることで完成度を高めていく．

1.2 問題提起

ユーザテストの段階では，ユーザ体験の品質を高めて付加価値を向上させるために，プロトタイプをユーザが試用した際の利用体験を得ることが必要となる[1].

利用体験は主に行動，思考，感情の三種類からなり，利用体験を得ることで，ユーザがプロトタイプを使用する際に感じる問題点・不満点を得ることができる．これらの問題点・不満点から，開発者やデザイナーはユーザの求める機能や性能を定義できるため，プロトタイプの改良に利用体験は欠かせない．しかし，その一方で，特に観察のみでは記録できない思考や感情の取得手法は定まっておらず，実際の開発現場で用いられている手法にも課題が存在する．

思考や感情といった，ユーザの内観を取得するために利用される情報が発話である．発話によって認知プロセスを分析する最も基本的な手法として，ユーザテスト後にインタビューを実施して振り返る手法が多く用いられている[2]．これは質問に対して，ユーザにテスト時の状況を振り返ってもらい，各行動時に何を考えていたか発話させる手法である．しかし，記憶によるバイアスや，状況を回顧する際に理由を後付けする恐れがある．

そのため，ユーザの内観をその場で得ることのできる調査手法が求められる．そこで利用されるのが，逐一考えていることを発話してもらう Think Aloud 法（以下 TA 法）である．Priede et al. [3]は，TA 法がインタビューに比べ，回答者の理解を調べることにより適していると述べている．しかし，ユーザは操作と並行して発話することになり，認知的負荷は高まる．よって，操作に真剣になるあまり発話を忘れてしまう恐れがある．これに対して，ユーザテスト中に他者の介入（質問や相槌）によって発話を促す事例も存在する [1][2]．しかし多くの場合，それらの介入は開発者やモデレータと呼ばれるユーザリサーチの専門家によって行われるため，ユーザの率直な発話が得られにくくなる．

そこで，本研究ではエージェントが介入する TA 法に着目し，ユーザテストへの応用の可能性を模索した．人間の代わりにエージェントが質問や相槌を行うことで，開発者には話しにくいと思われるプロトタイプへの不満や，インタフェース上の問題点を聞き出せるようになることが期待できる．ただし，プロトタイプとは独立にエージェントが存在してユーザに介入する場合，プロトタイプではなくエージェントに関心が集まってしまうことが考えられる．したがって，ユーザが抱くエージェントへの関心をプロトタイプへの関心に結びつけるため，プロトタイプそのものに身体的特徴を付与しエージェント化する．質問や相槌は音声で行い，プロトタイプが直接語りかけてくるように感じさせることで，通常の TA 法や開発者の介入と比較してよりユーザに話しやすい印象を与え，多くの発話を促す．その結果として，より率直な反応をユーザから引出すと同時に，ユーザビリティ改善のヒントとなる問題点への言及を得られる手法を目指す．

1.3 研究目的

本研究ではユーザテストを対象に，プロトタイプをエージェント化し，ユーザの発話を引き出す手法を提案する．

エージェント化したプロトタイプを利用することで，TA 法の最中にユーザが開発者等に抱く，発話に対する心理的障壁を軽減できるメリットが考えられる．これにより，ユーザの発話を増加させ，率直な感想を得られるようにすることを目的とする．また，分かりやすさや使いやすさといった，ユーザビリティへの言及に関しても同様に増加させることを第二の目的とした．これらの目的が達成されたかは，印象評価や発話内容の分析で確認する．

1.4 論文の構成

本論文の構成を以下に述べる．

第 1 章では，本研究の研究背景と目的について述べた．第 2 章では関連研究を紹介する．第 3 章では提案手法について説明し，第 4 章では，予備的実験と本実験について概要を述べる．第 5 章では実験結果と各種データに基づく考察を述べ，第 6 章では本研究のまとめと今後の課題を述べる．

第2章 関連研究

2.1 ユーザテストにおける TA 法の実施方法と介入の影響に関する研究

タスク中のユーザへの開発者の介入について，Boren et al.[4]は，連続的かつ控えめな肯定があると，タスクに集中したまま発話できるため，相槌やフォローアップクエスチョンが有効であると述べている．それに加え，予想外の応答で発話が止まった場合に質問することで，ユーザの説明を促せるとしている．一方で開発者の介入によって，ユーザは誤認識・誤操作の責任を自らに感じてしまう恐れがあることも指摘している．ユーザテストで発生する問題の多くは，テストされる製品に原因があり，ユーザの責任ではない．したがって，開発者はユーザと関係構築する上で，バグや予想外の出力結果の発生，操作方法がわからない等の問題発見はユーザの貢献であり，責任の所在は製品にあると強調するべきだと Boren et al.は結論づけている．

また，Olmsted-Hawala et al.[5]の実験では，TA 法を用いた web サイトのユーザテスト中にユーザに質問することで，タスクの成功率と web サイトへの満足度が上がることが指摘されている．この結果より，TA 法への開発者の介入によって，ユーザが抱くプロダクトへの印象は向上し，正当な評価でなくなることが考えられる．Krahmer et al. [6]は，厳密に定義された対話のない TA 法と確認応答のある TA 法とでユーザテストを実施し，被験者の評価や検出できた問題点の数に差がないことを示している．これらの研究から，TA 法に開発者が介入することで，タスクの成功率や満足度など，印象は本来の評価より高く評価され，責任の所在についても課題は残るが，ユーザビリティ上の問題点の検出精度への影響はないと考えられる．

しかし，これらはいずれも開発者による介入が TA 法に及ぼす影響についての研究である．すなわち，人間がユーザに対して質問や相槌する形式であり，その役割をエージェントが担った場合については分析や議論は行われてない．

2.2 プロダクトやアプリケーションのエージェント化に関する研究

DiSalvo et al.[7]の研究が示しているように，人間に似た特徴を持つ物体に対して，ユーザはよりインタラクションを行う傾向がある．そこで，プロダクトやアプリケーションに，目や手などの身体的特徴を付与することでエージェント化する先行研究が多く存在する．これらの手法は直接擬人化手法とも呼ばれ，大澤ら[8]は家電製品を題材として説明対象をエージェント化することで，ユーザが説明対象により注視しやすくなり，関心が対象に集まると述べている．Iwamoto et al.[9]は，店頭の商品紹介において，ヒューマノイドロボットが商品説明する条件と擬人化された商品が自己紹介する条件を比較した結果，商品説明をユーザが記憶しやすくなると結論づけている．さらに，自己紹介条件がヒューマノイドロボット条件と比べて，質問に対し否定する被験者の割合が多かった点から，被験者の率直な感想が得られる可能性が示唆された．

しかし，これらのエージェント化の事例の多くは店頭での商品説明などを想定しており，擬人化されたエージェントが自己紹介する，いわば情報伝達に特化した研究である．そのため，ユーザテストのように人間の発話から情報を引き出す目的には対応していない．本研究はユーザの発話を支援するために擬人化を用いる点が特徴的である．

第3章 提案手法

3.1 提案手法の概要

プロトタイプをエージェント化することで，ユーザテスト中にユーザとプロトタイプの直接対話を可能にすることを目指す．本研究では，プロトタイプのエージェント化研究の探索的な試みとして，対話機能を付加しやすい GUI 上のアプリケーションを対象に手法を設計する．

プロトタイプは TA 法に基づいたユーザの発話に対して，音声形式で応答・質問する．また，プロトタイプがユーザの話や操作を理解しようと努めていることをユーザに感じさせるため．画面上にマウスカーソルを追う目を付与する．以後，提案手法を構成する要素である，音声対話と身体的特徴として付与された目について，それぞれ説明する．

3.2 セリフの選定・音声合成

ユーザに，プロトタイプそのものが発話していると感じさせることを念頭に手法を設計した．まず，エージェント化したプロトタイプが話すセリフを用意するため，定型文を Coefont[10]で音声合成した．定型文の内容は表 3-1 に示すように，Boren et al.[4]の議論に基づき，ユーザへの教示から始まり，TA 法によるユーザの発話に対する相槌や，発話が止まった際に促す意図の内容に絞って生成した．音声はシステムらしさを考慮し，Amazon Alexa を参考に女性ナレーターの声を選択した．ユーザの発話内容に応じた定型文を再生し，ユーザとプロトタイプ間の対話を擬似的に実現する．実験中は Wizard of Oz 法（以下 WOZ 法）に基づいて実験者がシステムを装い，ユーザからは見えない PC の裏面に Bluetooth スピーカを配置し，ユーザの背後にいる実験者が，被験者に気づかれないように発話に合わせて，音声を選択・再生した．

表 3-1　定型文の内容

場面	セリフ
教示・自己紹介	1. 私はペイントソフトの（ソフトの名称）です。 よろしくお願いします！ 2. この目であなたの操作を見ています。 3. 私を使ってみて、あなたの感想を教えてください。
相槌	4. はい。 5. なるほど。 6. そうですね。
発話を促す質問	7. 今、何をしているところですか。 8. 何か悩んでいますか？ 9. どこで迷っていますか？ 10. 何かお探しですか？
意図を問う質問	11. それって何のことですか？ 12. それは何のために使うものですか？

3.3 身体的特徴の付与

また，プロトタイプに関心を持たせるため，身体的特徴を付与する．今回はWinEyes[11]を採用し，目をプロトタイプ上の画面上に配置した．WinEyes が起動している PC 画面上では，WinEyes の目玉がマウスカーソルの位置を追従する．この機能により，ユーザの操作を WinEyes が見つめているかのように思わせる効果が期待できる．実験でテストの対象とするアプリケーションに，実際に WinEyes を配置したスクリーンショットを図 3-1 に示す．

なお，これらの手法は，将来的にシステムとして自動化することを見据えたものである．そのため，発話のタイミングは秒数で決められている．具体的には，ユーザが 2 秒沈黙した場合，発話完了したとみなして相槌を打ち，10 秒沈黙した場合には現在の状況や悩んでいることを尋ねる．

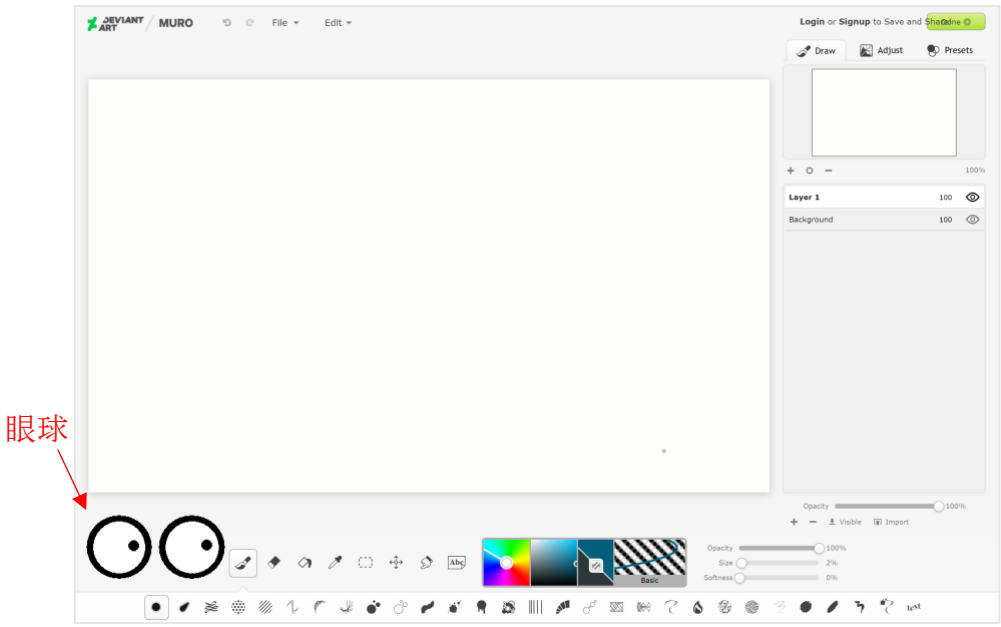


図 3-1　提案手法の画面イメージ

第4章 実験

プロトタイプのエージェント化による発話内容や参加態度の変化を検証するために行った実験について説明する．

4.1 実験計画

今回の実験は，各条件での印象の違いをユーザ役の被験者からインタビューで聞き出すために被験者内実験として実施した．また，ユーザテストの対象として web ブラウザ上で動作するペイントソフトを用いた．同じ操作感に慣れてしまうと，被験者が使い難い点やわかり難い点に直面しても言及せずにタスクを続ける恐れがあると考え，2 種類の異なるアプリケーションを用意した．今回の実験では，図 4-1 に示す通り，muro[12]と 8bit paint[13]の 2 種類を選定し，被験者に使用してもらった．この 2 つのソフトを利用することで，被験者は同等の難易度の 2 つのタスクに，それぞれ異なるインタフェースを使用して取り組むことになる．

被験者にはペイントソフトのユーザテストを行うと説明し，20 分間を目安に， 図 4-2 の 2 種類のイラストを条件別に用意し，できるだけ多くの機能を試しながら，イラストを正確に真似て描くように教示した．これらのイラストは同じタッチであり，描画の難易度は変わらないと判断した．また，被験者には，被験者の描く絵を評価するのではなく，ペイントソフトの使い勝手を評価することが目的であることを実験前に伝えている．

4.2 測定内容

今回の実験では，タスク中の被験者の操作画面を録画すると同時に，発話内容を録音した後に録音データを CLOVA Note[14]でテキスト化した．発話を定量的に扱うために，海保ら[15]の見解を参考に，文末を発話単位の終了と定め，言葉の途切れや言い淀みで切れ目が不明瞭な場合は 2 秒以上の間隔を発話単位の終了とみなし，1 分あ

たりの発話数（以下発話数）を集計した.

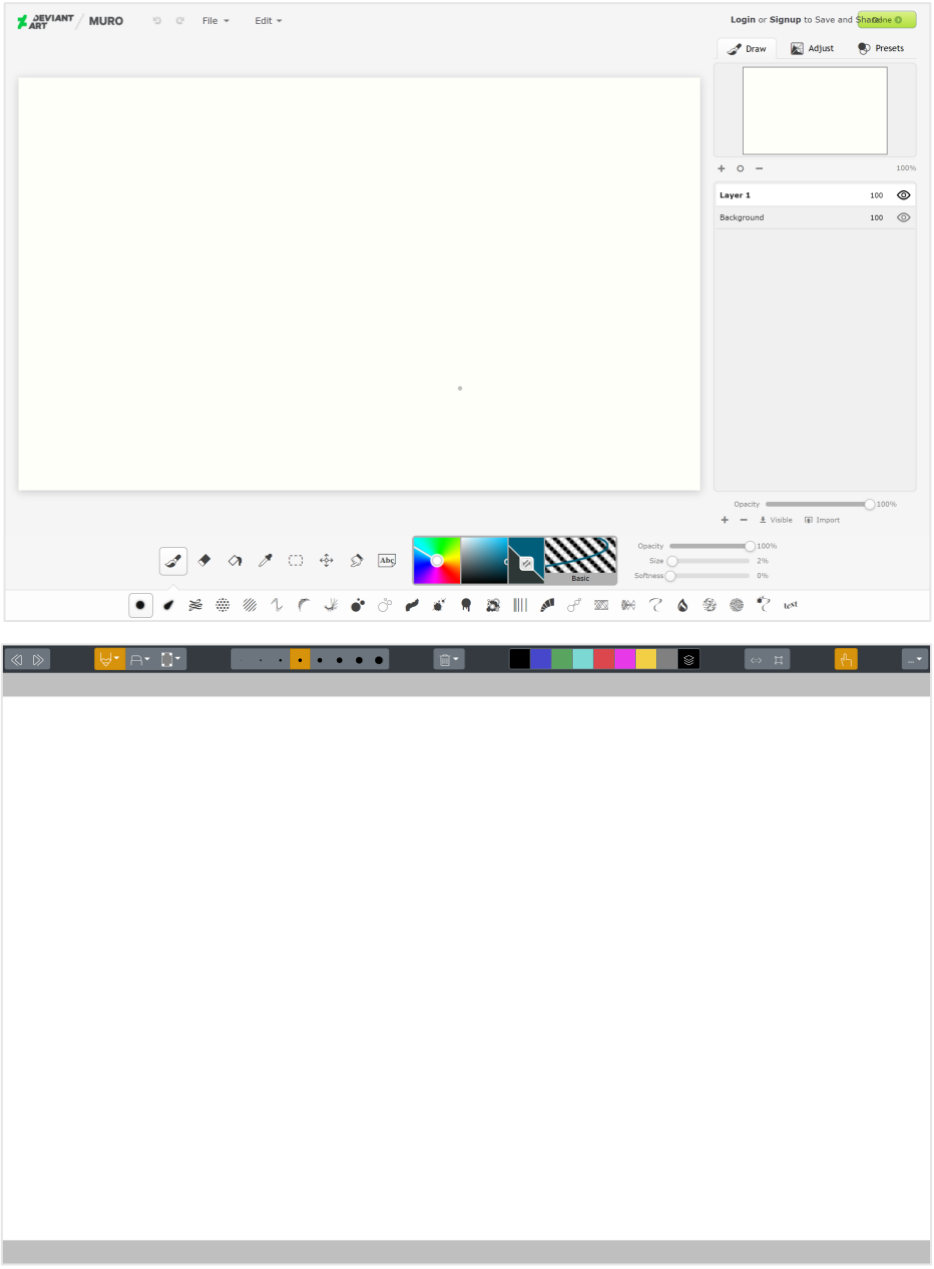


図 4-1　実験に用いたペイントソフト

（上：muro，下：8bit paint）

図 4-2　実験課題にしたイラスト

4.3 予備的実験

エージェント化したプロトタイプをユーザテストに用いることは，TA 法のセッション中，対話相手が人間からプロトタイプ本体に移行することを意味する．プロトタイプと対話することで，被験者の発話が増加すれば，エージェントのプロトタイプ化は被験者の発話を妨げることなく，ユーザテストを実施できるといえるだろう．しかし被験者の発話が減少する場合，提案手法の有効性が損なわれる可能性がある．

そこで，エージェント化されたプロトタイプを TA 法中に導入することによる，発話数への影響を分析するために，開発者が同席しない一般的な TA 法条件と提案手法を用いた Agent 条件の比較実験を行った．

TA 法条件はこの予備的実験の統制条件であり，実験中において被験者は何の反応も返ってこない状態で，TA 法による発話を行いながら操作する．Agent 条件は，提案手法に基づき，被験者の発話に合わせて実験者が WOZ 法で反応することで対話する．予備的実験での実験室内の状況を図 4-3 に示す．

被験者は本学の大学院生 2 名であり，対象とするペイントソフトを利用した経験がないことを確認した上で，実験に参加してもらった．なお，予備的実験に参加した 2 名は後述する本実験には参加していない．

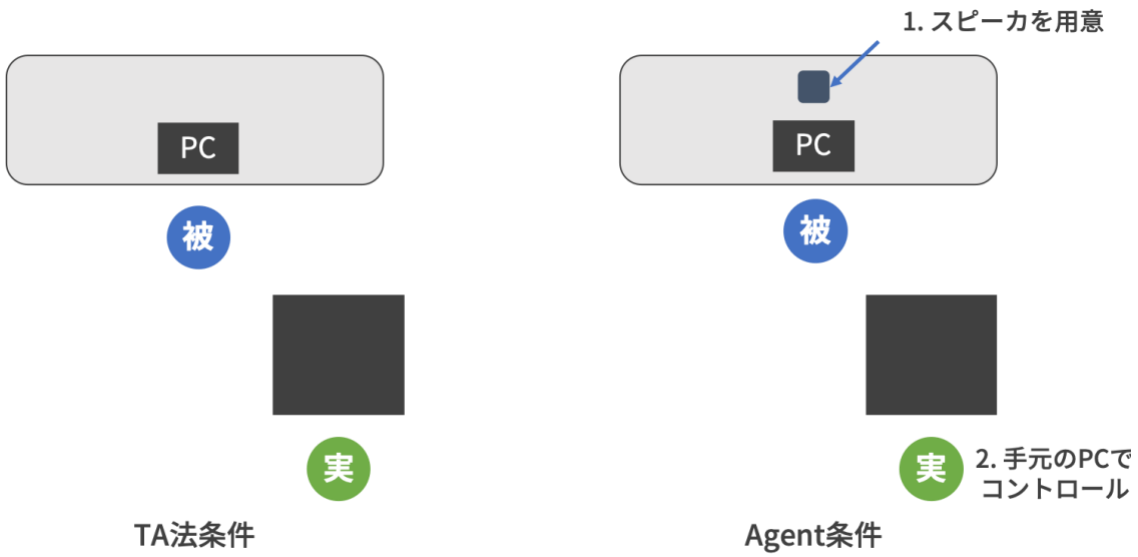


図 4-3　予備実験中の実験室内

（被：被験者，実：実験者）

予備的実験で得られた，被験者別の 1 分あたりの発話数を表 4-1 に示す．この結果より，予備的実験を通して，1 分あたりの発話数が増加し，エージェント化したプロトタイプが被験者の発話を阻害しないことを確認した．この結果より，TA 法の特徴である操作しながらの発話において，沈黙になってしまうという問題を解決できる可能性が示唆された．また，実験後のインタビューでは，「相槌を打ってくれる分だけ多く話したと思う（被験者 A)」という感想も被験者から引き出すことができた．

表 4-1　予備的実験での 1 分あたりの発話数

	被験者 A	被験者 B
TA 法条件	1.39	2.60
Agent 条件	2.96	3.22

4.4 本実験

本実験では，開発者が横につく Human 条件を統制条件とし，提案手法である Agent 条件との比較実験を行った．

Human 条件では，開発者と名乗る人間が被験者の隣に座り，被験者の発話に反応

することで対話する．その際の開発者の発話内容やタイミングは提案手法と同じ条件に揃えている．開発者役は本学の大学院生 1 名が行い，全被験者に対して同一の人間がついて実験した．Agent 条件では，予備的実験同様，被験者の発話に合わせて実験者が WOZ 法で反応することで対話した．実験室内の状況を図 4-4 に示す．

被験者は本学の大学院生 4 名であり，予備的実験と同様，対象とするペイントソフトを利用した経験がないことを確認した上で，実験に参加してもらった．また順序効果を避けるため，被験者ごとに各条件の実施する順番を入れ替え，表 4-2 に示すように，それぞれの順番ごとに同じ人数になるように調整した．なお，どの被験者も表の上から下に実施している．

予備的実験から本実験へ移行するにあたり，発話数に加えて，Norman [16]の提唱する行為の 7 段階に問題が生じていそうな場面を，被験者の発話から集計し，1 分あ

図 4-4　本実験中の実験室内

（被：被験者，実：実験者，開：開発者）

表 4-2　実験計画（H：Human 条件，A：Agent 条件）

	使用ソフト	被験者 A	被験者 B	被験者 C	被験者 D
1 回目	8bit paint	H	A	H	A
2 回目	muro	A	H	A	H

たりのユーザビリティ上の問題点への言及数（以下言及数）として記録している．行為の 7 段階とは, ユーザの行為を①ゴールの形成, ②行為のプラン, ③行為の詳細化, ④行為の実行，⑤外界の状況の知覚，⑥外界の状況の解釈，⑦結果の評価の 7 段階に分類できると Norman が提唱した理論であり，主に②③④が実行過程，⑤⑥⑦が評価過程とされる．本実験では，画面上のどこを操作すれば良いかわからない場合や，予想外の出力結果が生じた場合に実行・評価に支障をきたしていると筆者が判断し，そのアクシデントに対する被験者の発話を言及数として集計した．

また，Human / Agent の各条件のタスク後に，アンケートによる印象評価を行った．アンケートでは,「話しやすさ」と「関係性」の二つの観点で調査した．「話しやすさ」は，率直な発言ができたかどうか（以下率直性）を表す「考えたことや感じたことを素直に話せた」をはじめ,「開発者の機嫌が悪くなりそうなことも話した」などの 5 項目からなる．「関係性」は被験者と開発者，またはエージェント化したプロトタイプとの関係性を指し，ユーザビリティ上の問題発生時における責任の所在を問う「うまく操作できないときでもアプリ側の問題だと思えた」などの 5 項目を用意した．

TA 法を利用したユーザテストを実施する場合，開発者側には問題点そのものや，その特定に役立つような感想・意見をユーザの発話から引き出すことが求められる．そこで,「話しやすさ」「関係性」とは別に，ユーザテストの妥当性を問うために「アプリの問題点を全て言うことができた」という 1 項目を用意した．いずれの質問項目でも，被験者には 7 段階評価（1:全くそう思わない～7:非常にそう思う）で回答してもらった．加えて，実験終了後に各条件後のアンケート結果を見返しながらインタビューを行い，それぞれの印象と Human / Agent の 2 条件の回答の差の理由について尋ねた．図 4-5 に実験手順を示す．

大まかな所要時間は実験説明に 10 分，Human / Agent の各条件 20 分の計 40 分，アンケートも同様に各 10 分の計 20 分，インタビューに 20 分の計 90 分であった．

図 4-5　実験手順

第5章 実験結果と考察

実験中の風景を図 5-1 に示す．図 5-1 の上は Human 条件の実験風景であり，左に映るのが開発者役で, 右でマウスを握り, 操作しているのが被験者である. 下の Agent 条件では，画面左下に目玉を配置し，PC の裏にスピーカを隠した状態で音声を再生することでプロトタイプとの対話を擬似的に再現した.

実験では，図 5-2 に示す被験者の操作を記録した画面録画や，録音データをテキス

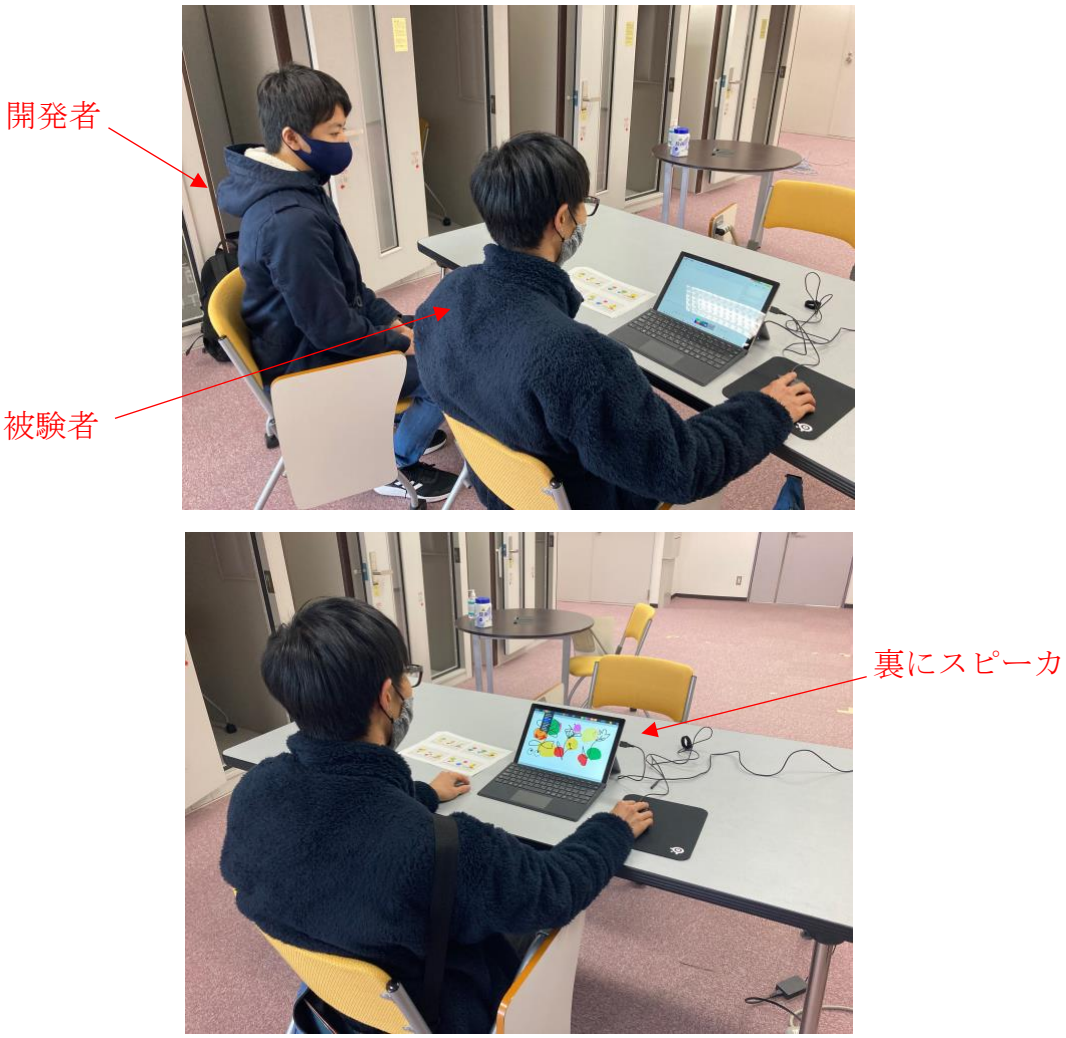


図 5-1　実験風景

（上：Human 条件，下：Agent 条件）

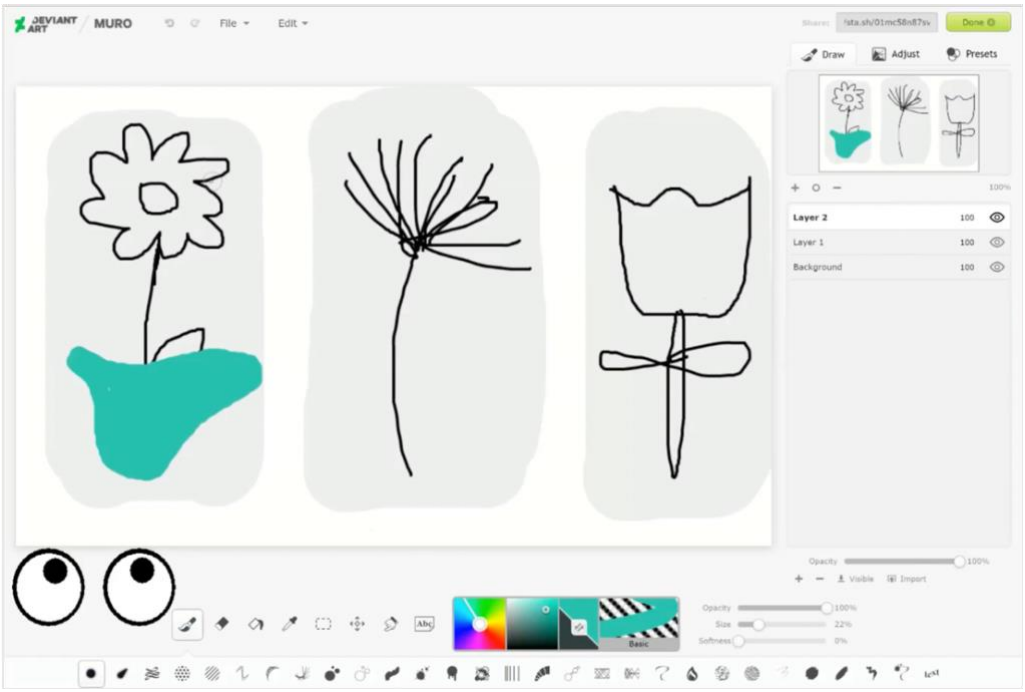


図 5-2　録画された実験中の操作状況（Agent 条件）

ト化して取得した発話内容，印象評価用アンケート及び実験後のインタビューなどのデータが得られた．そこで我々は，まずこれらの発話数・言及数やアンケート結果等の，定量的な結果の分析を行い，その後，発話内容やインタビュー内容に基づいてケーススタディ的に定性分析を行った．

5.1 アンケート・インタビューに基づく分析

印象評価用のアンケート結果に基づいて分析を行う．前述の通りアンケート項目は，被験者が開発者または，エージェント化されたプロトタイプに対して抱く印象を，「話しやすさ」「関係性」の二つの観点から評価するよう作成した．そこで，本節ではそれぞれの観点について分析し，インタビューで得た回答も振り返りながら考察する．なお，本実験では被験者 4 名と小規模であったため，アンケートの評価値による統計的仮説検定は行わなかった．

5.1.1 ユーザテスト中の「話しやすさ」の評価

「話しやすさ」の評価について，条件別に比較・検討した．「話しやすさ」の評価は多くの項目で一貫した傾向が見られず，個人差が大きかった．ただし，「話しやすさ」の中でも，率直性を表す「考えたことや感じたことを素直に話せた」の項目では，図 5-3 に示すように， 被験者 A，D の 4 人中 2 人が Agent 条件を Human 条件と比べて，高く評価した．インタビューでは，Human 条件については「悪いところを考えないようにしていた（被験者 C)」，「人間の方が具体的な返答をしてくれるのではと思っていた（被験者 A)」，「こうしたら伝わるかなと考えていた（被験者 B)」という感想が得られた．これより，Human 条件では，どのような内容を発話するかをより考えながら作業していたことが示唆された．その一方で Agent 条件では，「色んな意味で何を言ってもいいと思った（被験者 C)」との発言があった．また，各条件で同じ評価をした被験者 B も「同じことしか返さないのでわからないだろうなと思って話していた」と述べており，この反応より，定型文を再生するのみの Agent 条件に対して被験者は，何を言ってもいいと感じていたと考えられる．しかし，それは心理的障壁が軽減されたことによるものというより，反応の種類の少なさが影響しているとも考えられる．

ユーザの主観をそのまま伝えていたかを問う，「わがままだと思われそうな内容は思いついても話さなかった」の項目では，図 5-4 に示す通り，被験者 D のみ Agent 条件がよりわがままな内容を話したと回答した一方で，他の被験者はほぼ変わらない回

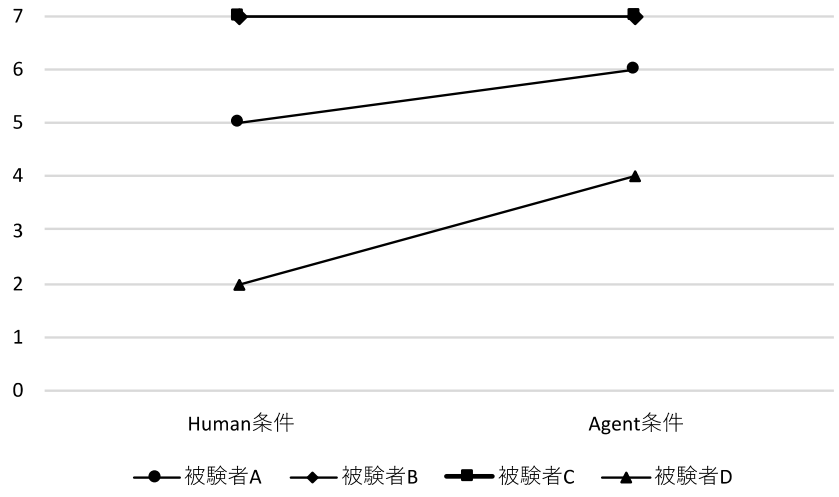


図 5-3 「考えたことや感じたことを素直に話せた」の被験者別評価

答結果となった.

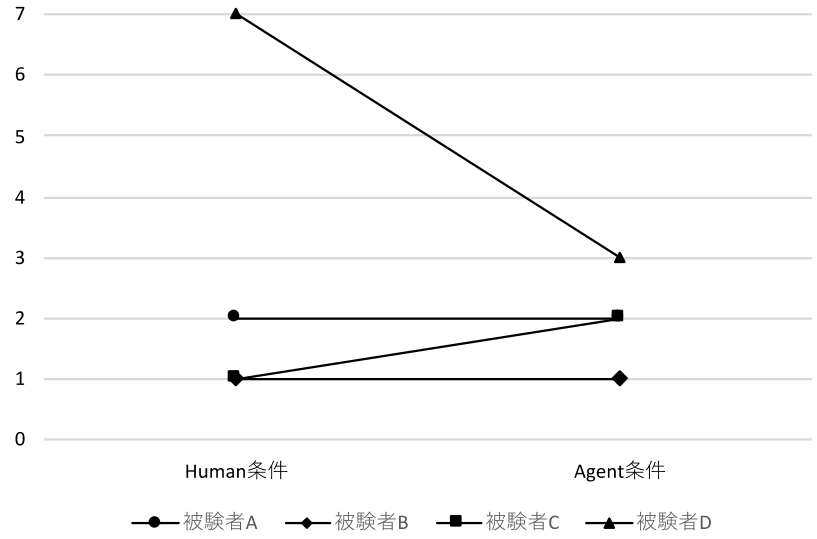


図 5-4 「わがままだと思われそうな内容は思いついても話さなかった」の被験者別評価

また，開発者・プロトタイプを気にせずに批判的な内容を話せていたかを問う「開発者（Agent 条件ではキャラクター）の機嫌が悪くなりそうな内容も話した」の項目では，図 5-5 に示す通り，被験者 A, D は Agent 条件の方を，被験者 B, C は Human 条件の方をそれぞれより話せたと評価した．被験者によって評価が分かれたが，被験者 B はインタビューで Agent 条件について「単純に疑問に思ったり、扱いづらいことを言った」と述べた一方，Human 条件については「開発したって仰ってたんで，ここが改善点だよと教える意味合いが強かった」と回答しており，批判的な内容を話せていたかの違いはみられなかったが，異なる内容を発話していたことが示唆された.

「開発者（Agent 条件ではキャラクター）の存在を気にせずに話した」では，図 5-6 に示す通り，被験者 4 人全員が Human 条件より Agent 条件を低く評価した．Human 条件の開発者に対して，Agent 条件のキャラクターをより気にしている可能性が示唆された.

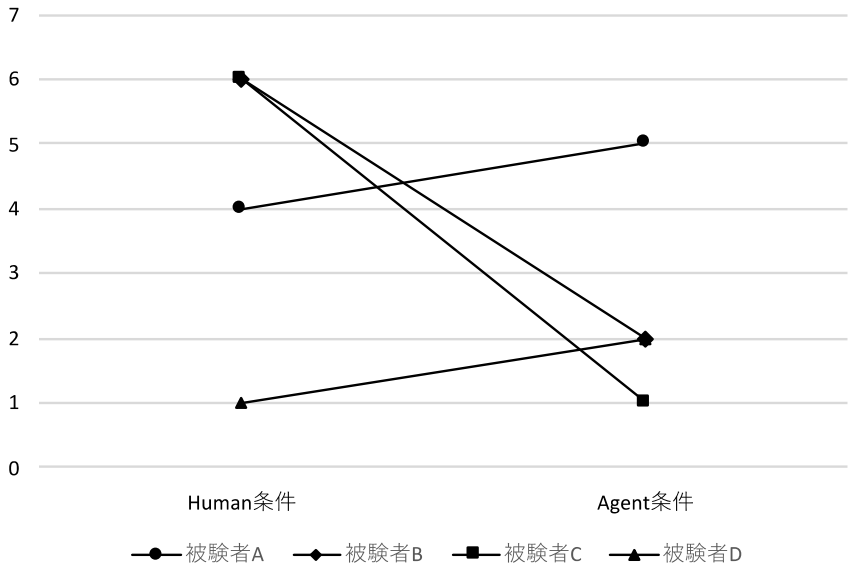


図 5-6　「開発者／キャラクターの機嫌が悪くなりそうな内容も話した」の被験者別評価

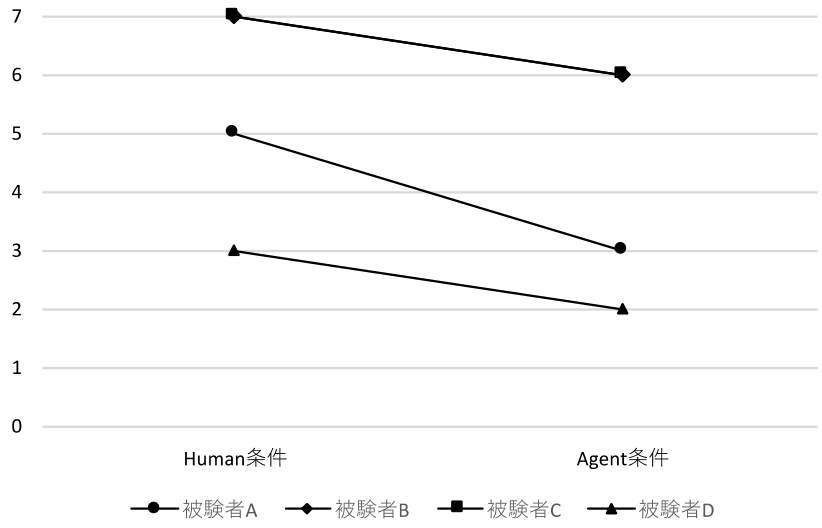


図 5-5　「開発者／キャラクターの存在を気にせずに話した」の被験者別評価

5.1.2 被験者との「関係性」の評価

実験中の被験者との「関係性」を 2 条件で比較すると，「話しやすさ」同様に，被験者によって各項目の評価は分かれる結果となった．しかし，「うまく操作できないときでもアプリ側の問題だと思えた」の項目では図 5-7 に示す通り，4 人中 3 人の被験者が Agent 条件を Human 条件と比べて高く評価した．これより，被験者がイメー

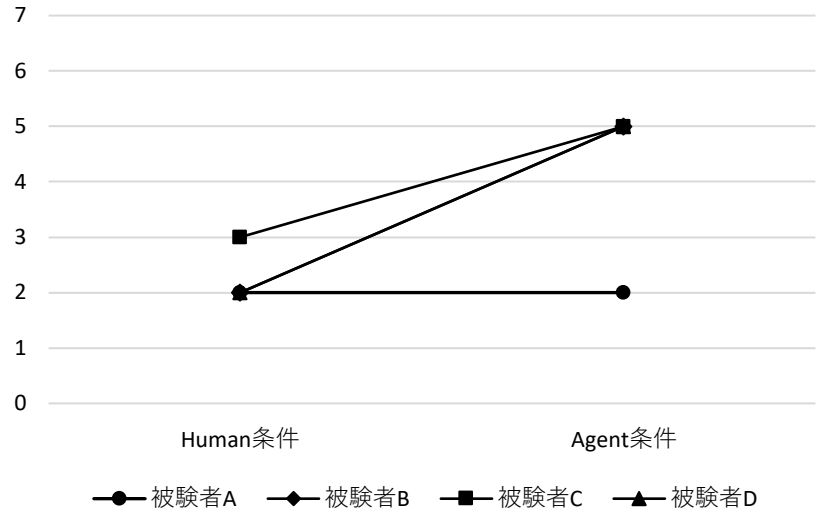


図 5-7 「うまく操作できないときでもアプリ側の問題だと思えた」の被験者別評価

ジ通りに操作できなかった際に， Agent 条件では Human 条件に比べて，ペイントソフト側にその責任があるとより感じていたことがわかった．インタビューでも，使用したペイントソフトの種類に，Agent 条件で扱った方のソフトを「アプリの表記に問題があった（被験者 B)」「必要なものが大体なく，能力の問題でない（被験者 C)」と評価していた．実際に，被験者 B と C は Agent 条件下で異なるソフトを操作している．以上により，プロトタイプをエージェント化することで，ユーザの予想の範疇にない出力結果になった場合に，その原因を自らの操作ではなく，ユーザテストの対象であるペイントソフトのユーザインタフェースに求めるようになることが考えられる．

また，責任の所在について，逆に自らにあると感じたかを「うまく操作できないとき自分のせいだと思った」という項目で尋ねた結果，図 5-8 に示す通り，被験者 A は同じ評価であったが，被験者 B, C, D は Agent 条件をより低く評価した．このことから，操作方法がわからないあるいは出力結果が予想通りでなかったときに，Human

条件の方が自らに責任を感じたことが考えられる．

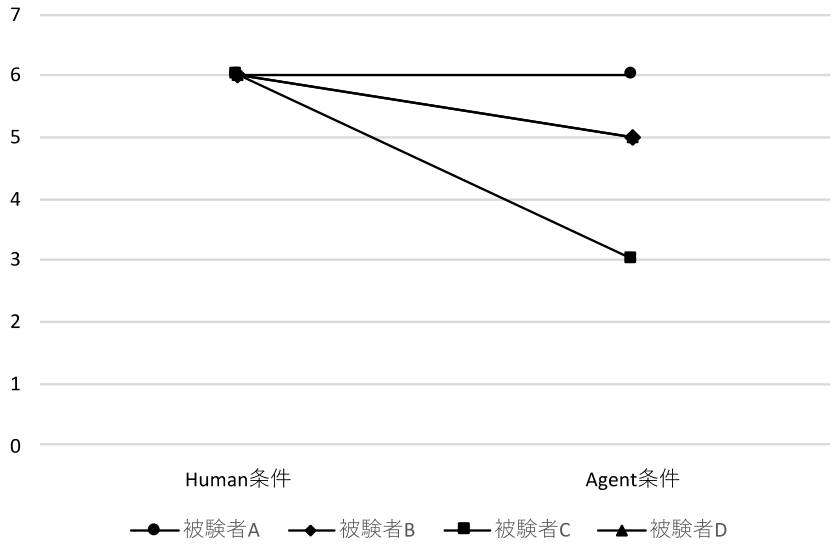


図 5-8 「うまく操作できないとき自分のせいだと思った」の被験者別評価

その他の項目では，緊張せずに発話できたかを問う「開発者（Agent 条件ではキャラクター）が喋ることで緊張した」では，図 5-9 に示すように，被験者 A, B は Agent 条件の方がより緊張したと評価し，被験者 C, D は Human 条件の方がより緊張したと評価した．被験者 A は「開発者が今回学生みたいな感じだったので，もしこの開発者が本当に知らない人とかなら変わると思います」と，開発者役が学生であったこと

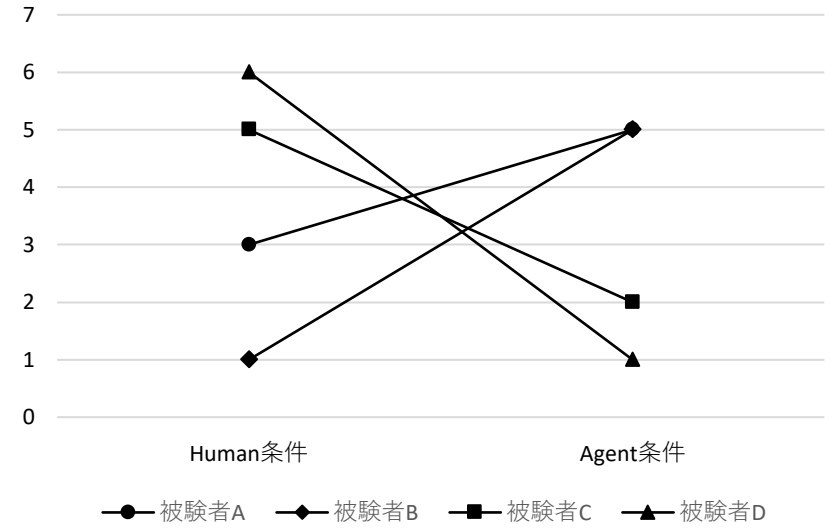


図 5-9 「開発者／キャラクターが喋ることで緊張した」の被験者別評価

に緊張の違いの原因があると指摘した．被験者 B は「(Agent 条件は）うまく伝わらないっていう緊張はあったかもしれない」と振り返っていた．これより，被験者が緊張せずに発話できるようにするには，エージェント側が被験者の話を聞きながら発話しているという姿勢を，被験者に示す必要があることが考えられる．

「自分と開発者（Agent 条件ではキャラクター）は対等な関係だと感じた」の項目では，図 5-10 に示す通り，被験者 A, D は Agent 条件をより高く評価し，被験者 B, C は Human 条件をより高く評価した．この結果より，ユーザテスト中に開発者が上の立場にいると考える傾向をプロトタイプのエージェント化で解消するには本実験の手法では不十分であったと考えられる. インタビューでは, Agent 条件について「対等かどうかを意識しなかった（被験者 A)」という感想が得られた．

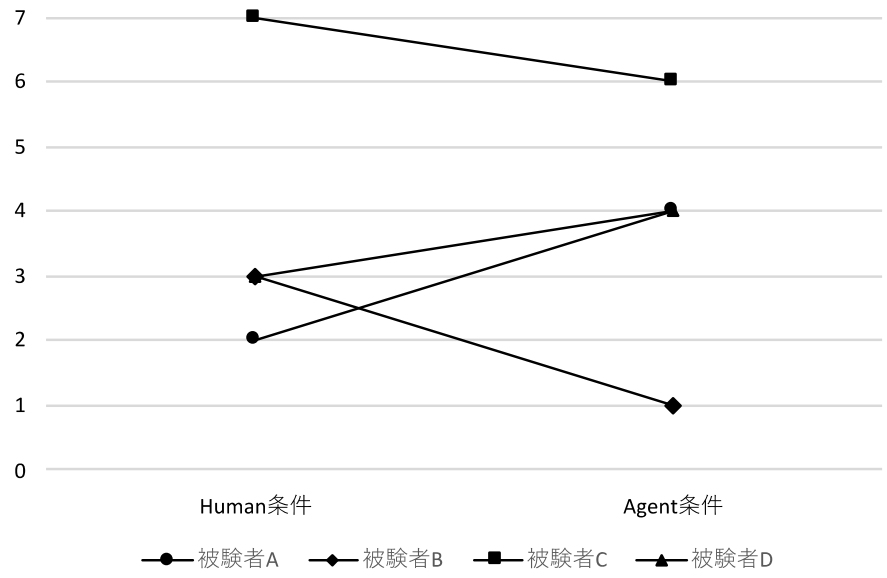


図 5-10　「開発者／キャラクターは対等な関係だと感じた」の被験者別評価

「開発者 (Agent 条件ではキャラクター) は何を言っても構わない相手だと感じた」の項目では，図 5-11 に示すように，被験者 A, C は Agent 条件を高く評価し，被験者 B, D は Human 条件を高く評価した．この結果は表 4-2 の実験計画が示す，先に実施した条件を低く，後から実施した条件を高く評価している．よって，実験中に TA 法

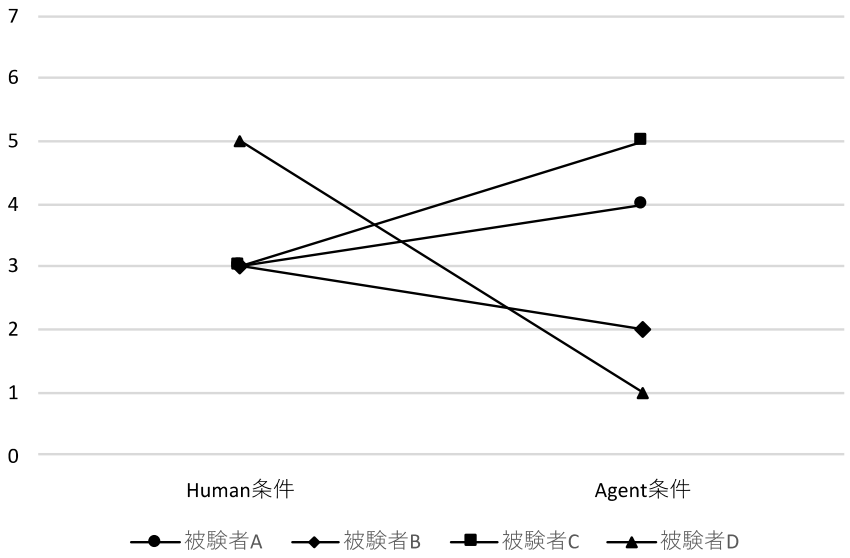


図 5-11 「開発者／キャラクターは何を言っても構わない相手だと感じた」の被験者別評価

に慣れてきたため，何を言っても構わないと感じるようになった可能性が考えられる．

5.1.3 ユーザテストに関する項目の評価

ユーザテストの妥当性を問う「アプリの問題点を全て言うことができた」という項目では，図 5-12 に示すように被験者 A, C は Agent 条件を高く評価し，被験者 B, D は Human 条件を高く評価した．被験者 A は，インタビューで Human 条件を 5，Agent 条件を 6 とそれぞれ評価し，その理由を「(Agent 条件の方が）質問頻度が高かったから」と振り返っていた．これには，被験者の機嫌に関係なく質問していたために，頻度が高くなったことが影響している．その一方で Agent 条件をより低く評価した被験者 B は，「補足説明をしたかどうかだと思う．キャラクターの方が問題点は言えなかった.」と話しており，どのような問題が発生しているかという詳細な説明

は，定型文しか返事しない Agent 条件には話しにくいと感じていたと考えられる．

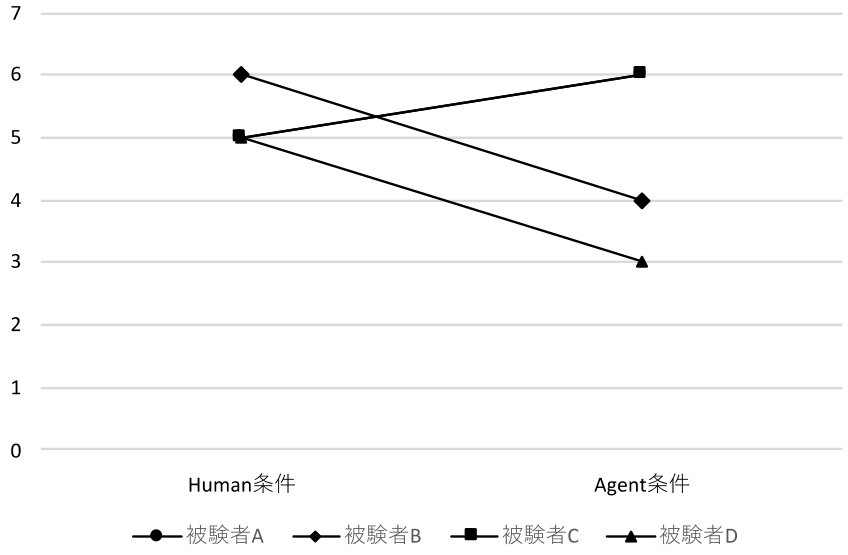


図 5-12 「アプリの問題点を全て言うことができた」の被験者別評価

5.1.4 アンケート・インタビューによる印象評価のまとめ

ここまでの内容を振り返ると，アンケート・インタビューによる印象評価は被験者によって異なり，「話しやすさ」と「関係性」のいずれにおいても，明確に片方の条件がより高く評価されたわけではない．ただし，各項目を個別に見ていくと多くの被験者が同じ傾向を示している項目も存在した．

「考えたことや感じたことを素直に話せた」の項目では，4 人中 2 人が Agent 条件を高く評価し，残り 2 人は同じ評価であったが，そのうちの 1 人のインタビューからも，Agent 条件におけるプロトタイプには何を言っても良いと感じていたことがわかった．

また，操作中に不明な点があった場合や，予想外の出力発生時の責任の所在について，被験者側の操作に原因があるとより感じるのは Human 条件であることがわかった．それに対し，プロダクト側の機能や画面配置に原因があると被験者がより感じたのは Agent 条件であることがわかった．インタビューからも，使用したペイントソフトの種類に拘らず，Agent 条件で扱った方のソフトに対し，機能に不備があると感じていたことが示唆された．このことより，プロダクトの改善を目的にユーザテストを

実施する際は，被験者が自らの責任にせず指摘する Agent 条件の方が，改善点をより多く発見できることが期待できる．

加えて，被験者の回答は 2 人ずつ分かれたが，提案手法を改良する上での課題が見えた項目もあった．「開発者（Agent 条件ではキャラクター）が喋ることで緊張した」では，Agent 条件ではうまく伝わっているか不安になるために喋ることに緊張したことがわかった．さらに，「アプリの問題点を全て言うことができた」では，Agent 条件では問題点に直面した際，使いにくいという感触は話せても，その詳細については話しにくかったことがインタビューでわかった．

このように被験者が緊張し，問題点を話しづらいと感じる原因として，定型文しか返ってこない点や，話を聞いているように思えなかった点が挙げられる．これらの解消には，話を聞いていることを被験者に伝える機能が必要となる．聞き手としての態度を示す方法には，相槌のバリエーション増加や相手の発話内容を引用しながら反応する等，対話中に示す方法と，身体的特徴である画面上の目を活かし，瞬きや視線の動きで示す方法の両面からアプローチすることが有効と考えられる．

5.2 実験中の発話数・言及数に基づく分析

テキスト化した発話データを元に集計し，定量的に分析する．発話数と言及数の算出過程は，予備実験同様に 4.2 節に示した基準に基づいて行った．その結果，表 5-1 に示すように，全体の発話数が全ての被験者で増え，表 5-2 に示す通り，言及数についても全ての被験者に増加がみられた．グラフを図 5-13 と図 5-14 にそれぞれ示す．

表 5-1　各被験者の 1 分あたりの発話数

	被験者 A	被験者 B	被験者 C	被験者 D
Human 条件	2.40	3.09	3.89	0.20
Agent 条件	4.75	4.56	4.76	1.48

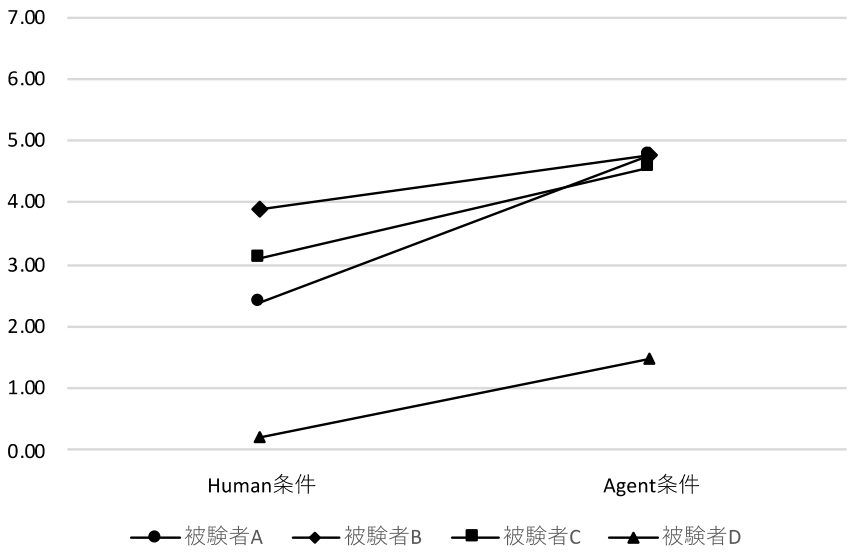


図 5-13　1 分あたり発話数の被験者別変化

この結果より，プロトタイプのエージェント化によって，ユーザテスト中における被験者の TA 法の発話が増加し，それに伴い，ユーザビリティ上の問題点に対する言及も増加したことが考えられる．ただし，Agent 条件は Human 条件に比べて相槌の頻度が増えたため, 発話が増加した可能性がある. また, 相槌増加の原因として, Agent 条件で WOZ 法を利用したことが挙げられる．教示があるとはいえ，どうしても開発者役が被験者の様子を伺ってしまう Human 条件に比べて，Agent 条件では被験者の機嫌を比較的伺うことなく相槌（音声再生）できるという性質を持つ．これにより，その影響を受けて相槌が増加したと推察できる．ただし，この特性は裏を返せば，被験者の機嫌に関係なく反応するエージェントの相槌が発話数を増加させうると考えることもできる．この知見は将来的に本手法を自動化した際に有用性が期待できる．

表 5-2　各被験者の 1 分あたりの言及数

	被験者 A	被験者 B	被験者 C	被験者 D
Human 条件	0.59	0.39	0.89	0.10
Agent 条件	1.05	1.07	1.35	0.14

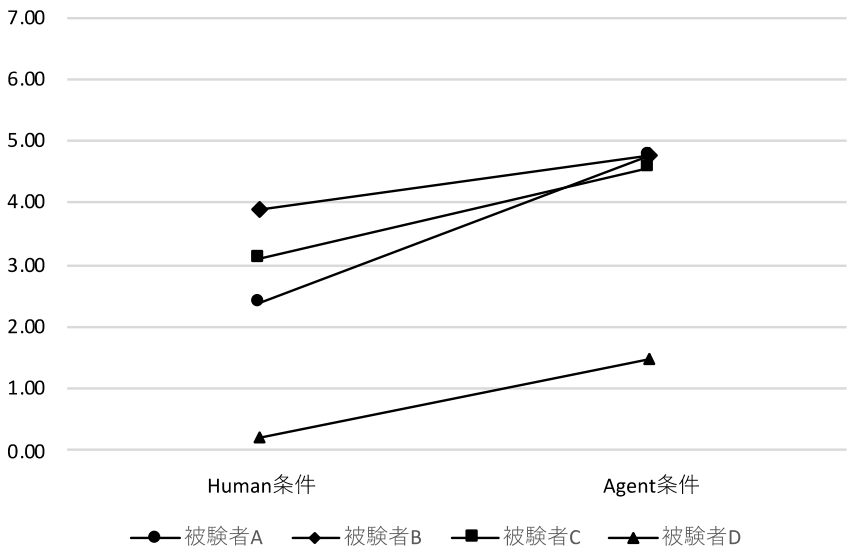


図 5-14　1 分あたり言及数の被験者別変化

5.3 発話データに基づく分析

実験で得られた発話データや画面録画の観察に基づき，各条件間の発話内容や表現の違いから，ケーススタディ的に分析・考察を行う．

5.3.1 全体的な口調の違い

被験者 A の口調が Agent 条件のみ徐々に敬語から言い放つような口調になっている．例えば，Agent の「何か悩んでいますか」という問いかけに対して，被験者 A は図 5-15 に示すように，「いや，特に悩んではないです，普通に書いています.」と返しており，返答がおざなりになっている．インタビューでは，被験者 A は「開発者の方（Human 条件）が、会話ができて，普通に喋れた感じがする」,「機械的な方（Agent 条件）が，話しても具体的な返答は見込めないかなと思った」と回答している．開発者が隣で返答する Human 条件に対し, Agent 条件では定型文しか返ってこないため，被験者は発話時にどう返答されるかを考える必要はない．そのため，言い放つような口調になったと推測できる．

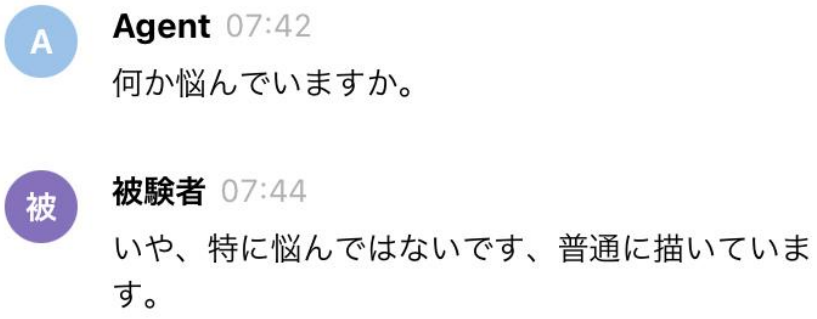


図 5-15　言い放つような口調の場面の例（被験者 A, Agent 条件）

5.3.2 操作の説明の違い

ユーザテストでは，ユーザの初見の感想を得るために，なるべく被験者には独力で操作することを促す [17]. すなわち長時間考えてもエラーから脱出できないなど，よほどの場合でない限り，正解を教えたり助けたりすることはない(本実験においても，答えを教えることはできないと教示している)．それにもかかわらず，図 5-16 に示すように，Human 条件では「とりあえずこれでいいのか（被験者 A)」,「ないものはないんですかね（被験者 D)」等，開発者に答えを求めるかのような口調が散見された．加えて，「普通に範囲を決めてあげて、そこをベタ塗りすればいいんですね（被験者 B)」のように，図 5-17 のように，自らの考えを開発者に確認しようとする場面もあ

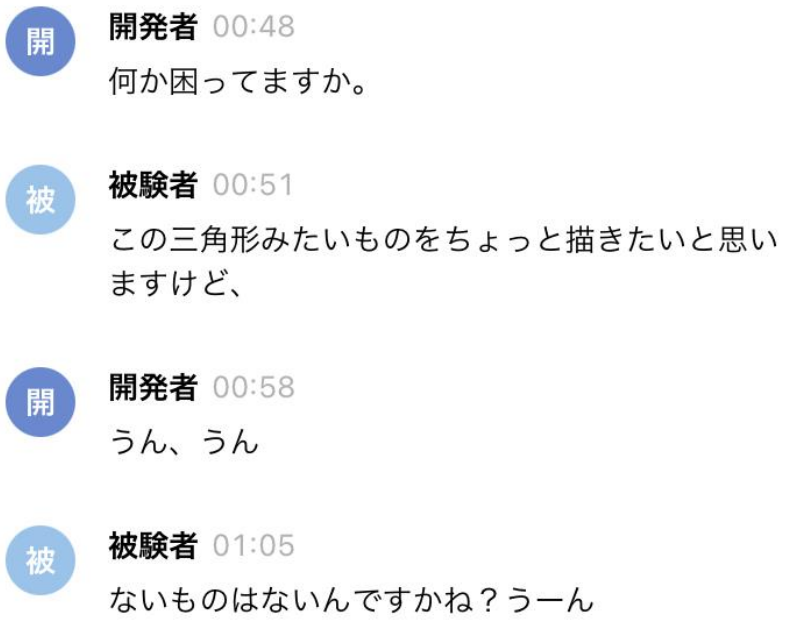


図 5-16　開発者に答えを求める場面の例（被験者 D, Human 条件）

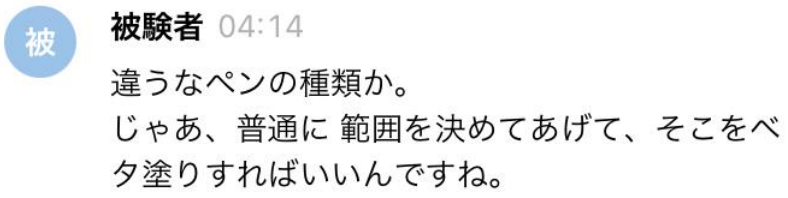


図 5-17　自らの考えを開発者に確認する場面の例（被験者 B, Human 条件）

った．

その一方で Agent 条件では，開発者不在のため，自分の考えを相手に説明するというよりも，図 5-18 のように，操作時の印象を直感のままに発話していた．実際，本実験においても，対話相手であるプロトタイプ側に答えを求めたり，確認したりする様子は見られなかった．人間の開発者と比べると，エージェント化プロトタイプは対話の相手とはみなされていないと思われる．よって，Agent 条件では，被験者は開発者の影響を一切受けておらず，その結果，「これで我慢するしかない（被験者 C)」に

図 5-18　操作内容をそのまま発話している場面の例（被験者 C, Agent 条件）

みられるように，被験者の率直な反応が発話に現れている．

その一方で，Human 条件では，図 5-19 に示すように，被験者は操作内容やその意図をより丁寧に説明していた．これは，開発者に対して伝えることを意識していたことが原因だと考えられる．

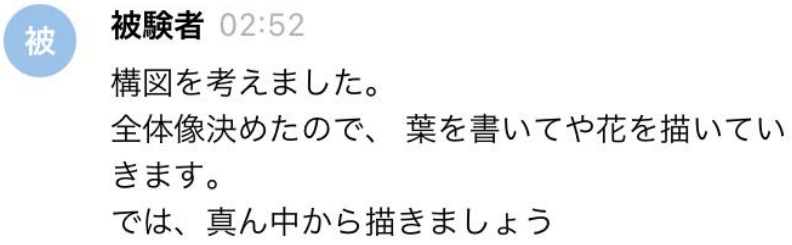


図 5-19　丁寧な操作説明の例（被験者 B, Human 条件）

5.3.3 責任の所在に関する表現の違い

被験者 B は図 5-20 に示す通り，操作の出力結果が予想と異なった際に，「色の重なりがやっぱり 1 番難しいですかね」等，「難しい」という表現を Human 条件では 7 回用いていた．「難しい」というフレーズから，思い通りに操作結果にならない原因をインタフェースだけでなく，自分の操作にも同程度求めていると筆者は推察した．また，Boren et al.[4]は，ユーザテスト中の開発者の介入によって，ユーザが誤操作を隠す恐れがあると指摘している．そのため，Human 条件においても，被験者は予想外の応答を誤操作であると感じ，「難しい」と言うことでそれを誤魔化していた可能性も考えられる．

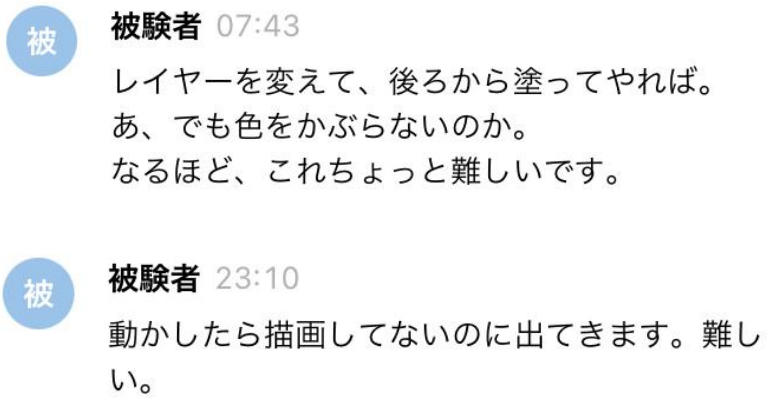


図 5-20　「難しい」と発話した場面の例（被験者 B, Human 条件）

それに対して，同じ被験者 B の Agent 条件では「難しい」は 1 回も出てこなかった．その代わりに，図 5-21 のように予想外の出力に対しては，使いたい機能がないアプリケーション側に問題があると判断し，そのことを指摘してすぐに次の作業に移る様子が窺えた．このように Agent 条件では，Human 条件のように婉曲な表現を使ったり誤魔化したりすることなく，アプリケーション側の問題点であると指摘していることがわかる．

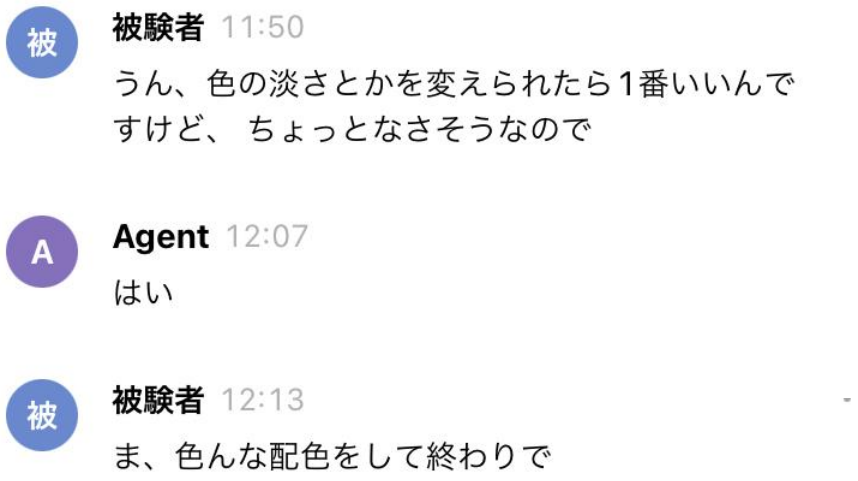


図 5-21　機能がないことを指摘する場面（被験者 B, Agent 条件）

5.3.4 ユーザビリティに関する言及・指摘時の違い

使いにくい点や分かりにくい点に直面したときの言及・指摘において，Human 条件では図 5-22 で示す「そのまま円という形状が出せるツールがないっていうのは，ちょっと不便かな（被験者 C)」のように，頭の中で状況や思考を整理し，何がどう不便なのか明確にしてから発話している．それに対し，Agent 条件では図 5-23 に示すように,「違う，こういうのじゃない（被験者 A)」,「これはどうするの（被験者 B)」,「アイコンが多すぎる、何が何なのかわからん（被験者 C)」等，操作する中で感じた違和感をそのまま言葉にしている．

筆者の推測ではあるが，これは 5.3.1 でも触れた通り，Human 条件では自分の考えを隣に座る開発者に伝えることを意識しており，一方 Agent 条件ではプロトタイプは考えを伝える対象として見做されていなかったためではないだろうか．

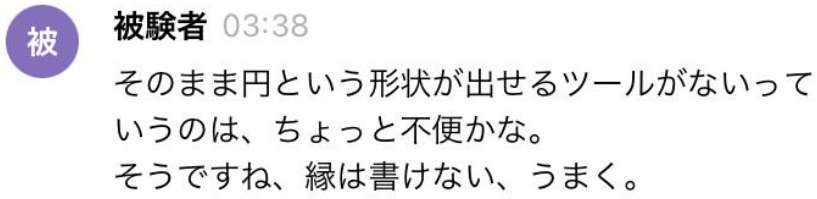


図 5-23　状況を整理してから発話している場面の例（被験者 C, Human 条件）

しかし，説明しようとする意識が生まれないことによるメリットもある．ユーザが言語化できない課題に直面した場合，Human 条件では説明を諦めてしまう恐れがあるが，Agent 条件では図 5-23 のように，違和感を感じたままに発話している．このことから，Agent 条件では，発話データから問題の箇所を特定するのが難しい一方で，その問題の存在を直感的な言及から発見できるメリットもあるかもしれない．ただしこれを活かすには，断片的な発話内容から課題のある機能を特定できるよう，具体的な状況把握が求められる．提案手法を実際の現場で応用する際には，発話データだけでなく，画面録画や視線情報などの記録を取る必要がある．

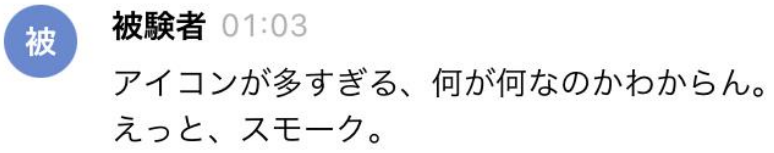


図 5-22　違和感を感じたままに発話している場面の例（被験者 C, Agent 条件）

5.3.5 Human 条件特有の機能を褒める発言

Human 条件では「ブラシの種類が多いのは面白い（被験者 B)」,「これは超面白いな（被験者 C)」等，図 5-24 に示す通り，Agent 条件では見られなかったペイントソ

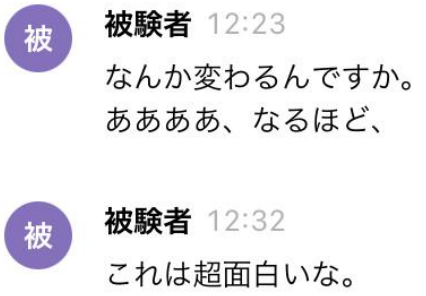


図 5-24　機能を褒める場面の例（被験者 C, Human 条件）

フトの機能を褒める発言があった．

また，図 5-25 に示す通り，被験者 B は Human 条件で「結構用途が特定されそうな筆の形ばっかりなので、あの、もうちょっとポピュラーのやつ増やしてもいいかもしれない」と指摘した後で,「使い勝手より面白みを追求してるような筆が多いです」とポジティブな発言をしていた．筆者の推察に過ぎないが，これは隣にいる開発者を気遣った結果，指摘だけで終わらせずにフォローを行った可能性がある．

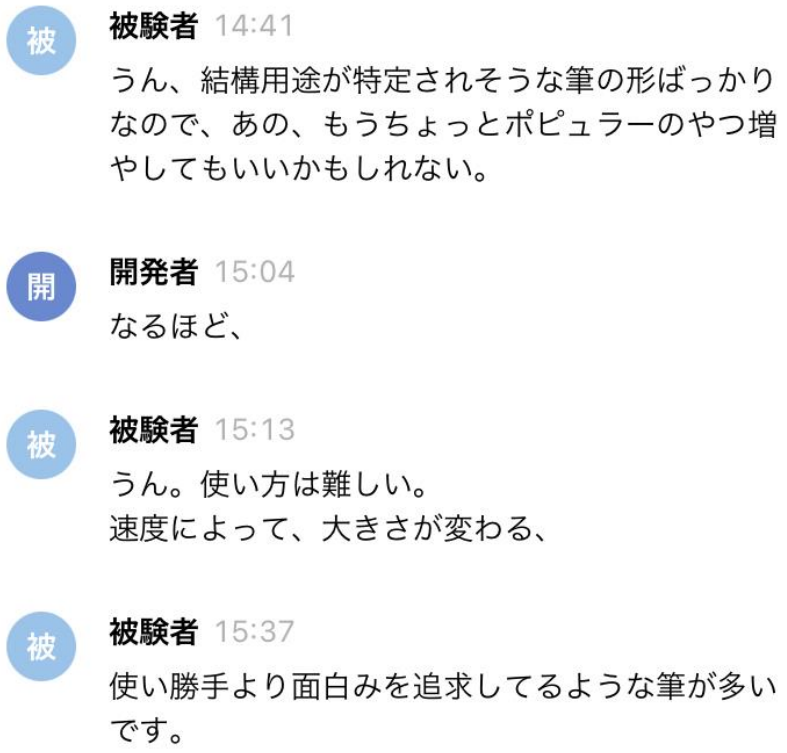


図 5-25　開発者に指摘した直後にフォローしている場面の例（被験者 B, Human 条件）

5.3.6 各条件の発話内容のまとめ

5.3.1 節から 5.3.5 節で述べた各条件の発話の特徴を表 5-3 に示すとともに，ユーザビリティ上の問題発生から発話までの一連の過程を表 5 を基に推測する． 被験者の対話相手が開発者とエージェント化プロトタイプとで異なることで，発話内容に大きな違いがあることがわかる．中でも「操作の説明」や「ユーザビリティへの言及・指摘」から，Agent 条件のプロトタイプは被験者にとって，対話相手と見做されておらず，加えて，Human 条件にみられるような，上手く説明しようという意識は生まれなかったことがわかる．

ここで，使いたい機能が見つからない場合や，予想外の出力結果などの問題発生から，発話するまでの一連の過程を表 5-3 を元に推測していく．まず，問題発生時は「問題発生時の責任の所在」から，Human 条件では自らの操作にも責任があると感じるのに対し，Agent 条件では機能に問題があると判断し，真っ先に指摘しようとする．

次に，その後の発話内容について，「ユーザビリティへの言及・指摘」から Human 条件では，開発者に伝えようと何がどう不便なのか明確にしてから，整理された明確な指摘を発話していた．それに対し Agent 条件では，操作中に感じた違和感をそのまま言語化していた．上手く説明しようという意識がないために，説明を諦めてしまうような抽象的な感覚も，Agent 条件では感じたままに発話していたと推察できる．以上を踏まえ，問題点発見から発話までの一連の流れを推測し，概念図にしたものを図 5-26 に示す．

上手く説明しようとするあまり，明確にできない問題点への言及を諦めている場面は言及していない時点で発話データには現れることはない．そのため，アンケートやインタビューを通して聞き出す必要がある．事前に設定していたアンケート項目・インタビュー内容の都合上，本実験では把握しきれなかった．しかし，言及や指摘を諦めることがあるかについて調査を進めることで，ユーザテスト中の TA 法における，ユーザの問題点発見から発話までのプロセスの判明にも役立てることができるのではないだろうか．

表 5-3　各条件の発話の特徴

	Human 条件	Agent 条件
対話相手	開発者	プロトタイプ
全体的な口調	• 常にです・ます調の丁寧語	• 最初は丁寧語だが，慣れてくると同時に言い放つような口調に
操作の説明	• 答えを求めて開発者に質問する • 自らの考えを開発者に確認しようとする • 開発者に対して伝えることを意識した，より丁寧に説明	• 操作内容や，それに付随する印象や感想をそのまま発話 • プロトタイプに返答はするが，対話相手だと見做されていない
問題発生時の責任の所在	• 思い通りの結果にならないのは，自らの操作にも責任があると判断	• 思い通りの結果にならないのは，プロダクト側に使いたい機能がないためと判断し指摘
ユーザビリティへの言及・指摘	• 頭の中で状況や思考を整理してから発話 • 何がどう不便なのか明確にし，開発者に伝わるように工夫している	• 操作中に感じた違和感をそのまま発話 • 印象や感覚についての言及が中心 • 被験者が上手く説明できない問題も，発話を諦めることなく，感じたままに発話している
その他特徴	• 機能を褒める発言 • 改善点を意見したあとにポジティブな発言でフォローする	

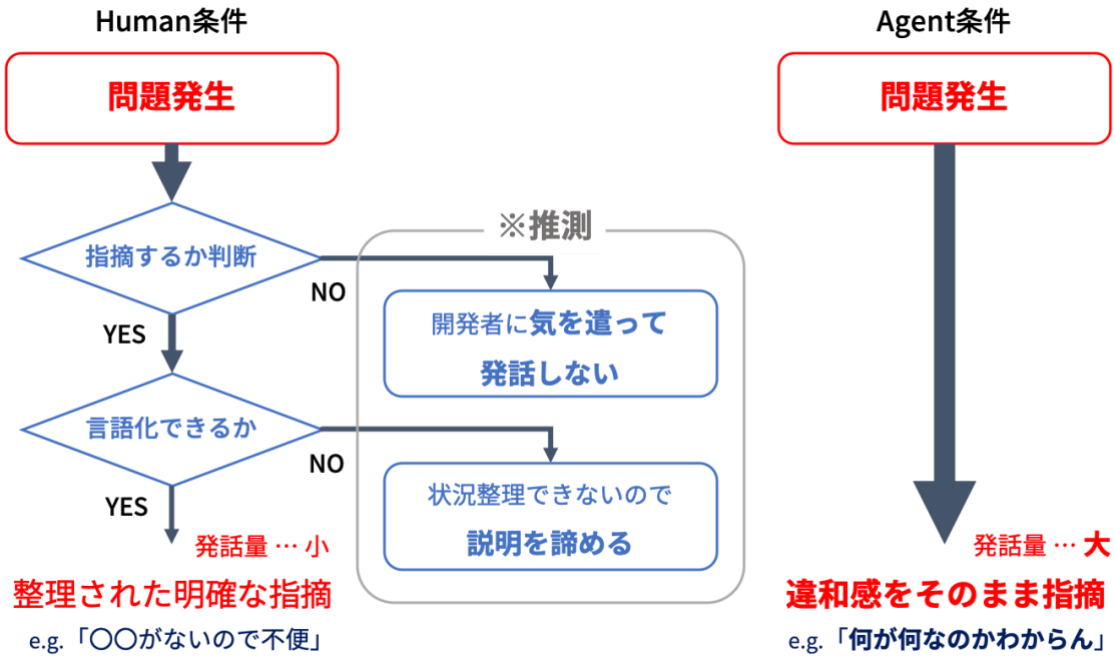


図 5-26　問題発生から発話までのプロセスの概要

第6章 結論

6.1 研究目的に対する回答

本研究ではユーザテストを対象に，エージェント化プロトタイプを TA 法の補助に利用することで，ユーザの発話やユーザビリティへの言及の増加と，正直な感想を得ることを目的とした．開発者が介入する Human 条件との比較実験の結果，全体の発話とユーザビリティへの言及のいずれも増加が確認された．また，発話内容の分析の結果，Agent 条件の方がより操作中に感じた違和感をそのまま発話していたことから，より正直な感想を得ることができたといえる．ただし，これらの発話数・言及数の増加や発話内容の変化は，研究当初に想定していた，開発者不在でプロトタイプが相手になることによる心理的障壁の軽減が原因というよりは，プロトタイプに対して説明しようという意識が，開発者相手と比較してより希薄であったことが影響していると考えられる．

ここで，1.1 節で述べた従来手法に対する，本研究の提案手法の有用性や立ち位置を再確認する．TA 法やその解決策として多く用いられる開発者の介入では，ユーザは開発者の存在を意識する分，熟考された内容が発話されるのに対し，プロトタイプ相手ではより直感に基づいた内容が得られる．そのため，従来の TA 法や開発者の介入に比べ，発話が増加した．よって，TA 法が持つ課題である認知的負荷の高さによる発話不足は，本研究の提案手法で補完でき，発話不足解消の観点では開発者の介入以上に効果的であると考えられる．

6.2 提案手法の有用性と各条件の使い分けの提案

本実験の結果から，プロダクトの改良を利用目的とした際の，提案手法の有用性を改めて述べる．Agent 条件の特徴である，問題発生時にプロダクト側に責任があるとユーザに感じさせる点は，より多くの指摘を得たい場合に有効であると考えられる．もう一つの特徴である，操作中の違和感を無理に状況整理せずに率直に発話する点は，プロダクトを利用する間のユーザが抱く感情推移の考察に役立つと考えられる．

よって，多少抽象的になったとしても直感的な反応や印象を得たい場面でエージェント化プロトタイプは特に効果を発揮すると考えられる．例えば，ユーザが操作する中でストレスを抱え得る箇所を，細部に至るまで洗い出したい場合には，エージェント化プロトタイプを用いると，違和感を逃さずに発話することから，より多くの改善点の発見が期待できる．

その一方で，従来の開発者が介入する手法である Human 条件にもメリットは存在する．特に，発話の絶対量が減少してでもユーザの意見や見解が欲しい場合には，対話内容や発話のタイミング等の精度にもよるが，人間の開発者が介入した方がより詳細な発言が得られると考えられる．機能に対する具体的なアドバイスを引き出して，ユーザと改善の可能性について議論する場合には，開発者による介入の上で実施することが望ましい．

このようにユーザ体験の調査精度の向上のためには，それぞれの特徴を理解した上で，開発者による介入とエージェント化プロトタイプとを目的に応じて使い分けることが望ましいと考えられる．

6.3 本研究のまとめ

本研究ではユーザテストを題材として，プロトタイプをエージェント化し，ユーザに直接話しかけることでユーザテスト中の TA 法を支援する手法を提案し，印象評価や発話集計，発話データの定性分析を通してその有用性を検証した．

「話しやすさ」や「関係性」などエージェント化されたプロトタイプに対する，被験者の主観的な印象評価は個人差の大きい結果となった．しかし，被験者の発話数を比較すると，Agent 条件では開発者が隣に座る Human 条件に比べて多く相槌した分，被験者の発話が増加した．その結果，ユーザビリティへの言及が促され，アプリケーションの機能や操作に対する違和感をより発話しやすくなることが示唆される結果となった．加えて，相槌の増加は被験者の機嫌に関係なく反応できるためであり，Agent 条件特有のメリットであると考えられる．また，操作方法がわからない場合や，予想外の出力があった場合等の問題発生時の責任の所在についても，Agent 条件ではユーザは自らの操作よりインタフェース側に原因があると考えることがわかった．発

話内容からも，Agent 条件の方がより抽象的な内容の発話が多く，筋道を立てたり含みを持たせたりせずに，その場ですぐに率直な反応を話す様子が窺えた．ただし，Human 条件にも，操作の説明をより丁寧に行う場面や，アプリケーションの機能を褒めながらアドバイスする様子が見られたため，今後は，実際の開発現場に応用する際はそれぞれを使い分けることが肝要である．

6.4 今後の展望

本研究は自動化を最終目標としているものの，まだ初期段階にあり，ユーザの率直な意見や改良につながるユーザビリティに対する指摘を最も得ることのできるエージェント化手法については十分に検討しきれていない．今後は，直感に基づく感想だけでなく，熟考された意見も開発者の介入と同じ，もしくはそれ以上に引き出すことが求められる．そのためには，返答のバリエーションの増加や，視線が動くあるいは頷きや瞬きなど，発話に対する反応を示す動作をエージェントに追加することが必要となる．これらの改良を通して緊張感の軽減，ひいてはユーザに「自分の話を聞いてくれている」と思わせることを目指す．

また，エージェント化プロトタイプとのインタラクションを設計する上で，よりユーザの意見を引き出すためには，どのような印象を与えるとよいか明らかにする必要がある．そこで有用と考えられる手法の例として，蔵田らの Talking-Ally に代表される弱いロボット[18]のように，プロトタイプ及びエージェントを，ユーザにとってより親近感を抱くことができる存在に改良した上で，同様の実験を進めることが必要である．

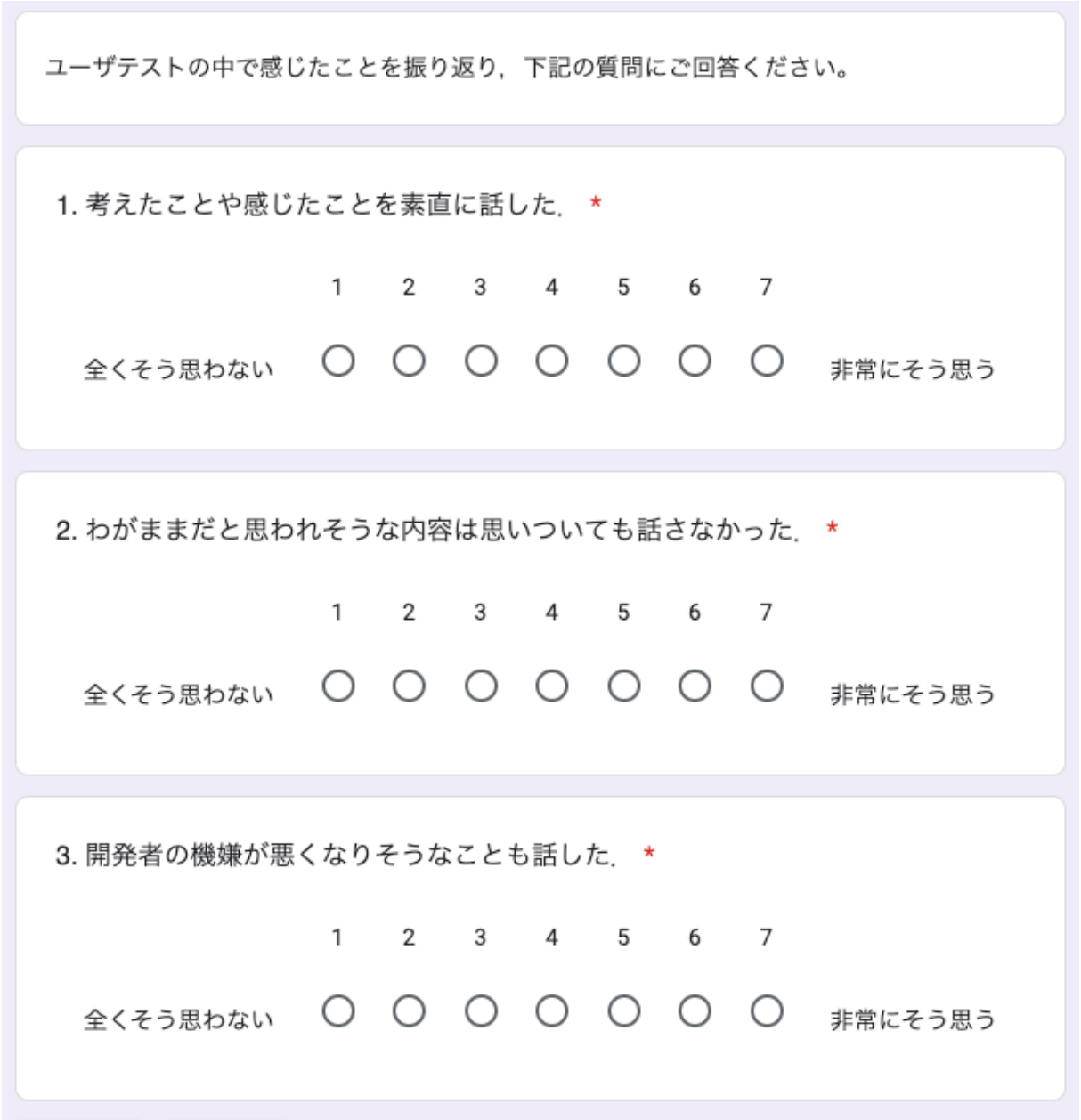
謝辞

本研究を進めるにあたり，研究課題の設定や研究に対する姿勢，本報告書の作成に至るまで，2 年間に渡って主指導教員として，全ての面で丁寧かつ熱心なご指導をいただきました，高島健太郎講師に厚く御礼申し上げます．また，ゼミでの進捗報告を通しての提案手法による課題解決に関する議論や，実験設計，測定内容並びに考察の方法など，終始貴重な助言を賜りました西本一志教授に心より感謝申し上げます．その他にも研究の遂行にあたり，多くの先生方からご支援をいただいたこと，改めて感謝申し上げます．

本実験で実験補助を担当してくださった西本研究室 堀江歩氏をはじめ，予備的実験から本実験に至るまでに協力してくださった全ての被験者の方々や研究室の皆様，そして，日頃から学生生活を通じて筆者の支えとなった友人にお礼申し上げます．最後に，高専卒業後も大学・大学院と勉学を継続することを快く受け入れ，応援してくれた両親に心から感謝いたします．

付録

6.5 Human 条件のアンケート



ユーザテストの中で感じたことを振り返り，下記の質問にご回答ください。

1. 考えたことや感じたことを素直に話した． *

	1	2	3	4	5	6	7	
全くそう思わない	○	○	○	○	○	○	○	非常にそう思う

2. わがままだと思われそうな内容は思いついても話さなかった． *

	1	2	3	4	5	6	7	
全くそう思わない	○	○	○	○	○	○	○	非常にそう思う

3. 開発者の機嫌が悪くなりそうなことも話した． *

	1	2	3	4	5	6	7	
全くそう思わない	○	○	○	○	○	○	○	非常にそう思う

4. 開発者の存在を気にせずに話した. *

	1	2	3	4	5	6	7	
全くそう思わない	○	○	○	○	○	○	○	非常にそう思う

5. 開発者がいることで緊張した. *

	1	2	3	4	5	6	7	
全くそう思わない	○	○	○	○	○	○	○	非常にそう思う

6. うまく操作できないときでもアプリ側の問題だと思えた *

	1	2	3	4	5	6	7	
全くそう思わない	○	○	○	○	○	○	○	非常にそう思う

7. うまく操作できないとき自分のせいだと思った *

	1	2	3	4	5	6	7	
全くそう思わない	○	○	○	○	○	○	○	非常にそう思う

8. 自分と開発者は対等な関係だと感じた. *

	1	2	3	4	5	6	7	
全くそう思わない	○	○	○	○	○	○	○	非常にそう思う

9. 開発者は何を言っても構わない相手だと感じた. *

	1	2	3	4	5	6	7	
全くそう思わない	○	○	○	○	○	○	○	非常にそう思う

10. アプリの問題点を全て言うことができた. *

	1	2	3	4	5	6	7	
全くそう思わない	○	○	○	○	○	○	○	非常にそう思う

6.6 Agent 条件のアンケート

1. 考えたことや感じたことを素直に話した. *

	1	2	3	4	5	6	7	
全くそう思わない	○	○	○	○	○	○	○	非常にそう思う

2. わがままだと思われそうな内容は思いついても話さなかった *

	1	2	3	4	5	6	7	
全くそう思わない	○	○	○	○	○	○	○	非常にそう思う

3. キャラクターの機嫌が悪くなりそうな内容も話した. *

	1	2	3	4	5	6	7	
全くそう思わない	○	○	○	○	○	○	○	非常にそう思う

4. 開発者の機嫌が悪くなりそうなこともキャラクターに話した. *

	1	2	3	4	5	6	7	
全くそう思わない	○	○	○	○	○	○	○	非常にそう思う

5. キャラクターの存在を気にせずに話した. *

	1	2	3	4	5	6	7	
全くそう思わない	○	○	○	○	○	○	○	非常にそう思う

6. キャラクターが喋ることで緊張した. *

	1	2	3	4	5	6	7	
全くそう思わない	○	○	○	○	○	○	○	非常にそう思う

7. うまく操作できないときでもアプリ側の問題だと思えた *

	1	2	3	4	5	6	7	
全くそう思わない	○	○	○	○	○	○	○	非常にそう思う

8. うまく操作できないとき自分のせいだと思った *

	1	2	3	4	5	6	7	
全くそう思わない	○	○	○	○	○	○	○	非常にそう思う

9. 自分とキャラクターは対等な関係だと感じた。 *

	1	2	3	4	5	6	7	
全くそう思わない	○	○	○	○	○	○	○	非常にそう思う

10. キャラクターは何を言っても構わない相手だと感じた。 *

	1	2	3	4	5	6	7	
全くそう思わない	○	○	○	○	○	○	○	非常にそう思う

11. 本当にアプリがキャラクターとしてしゃべっているように感じた *

	1	2	3	4	5	6	7	
全くそう思わない	○	○	○	○	○	○	○	非常にそう思う

12. アプリの問題点を全て言うことができた。 *

	1	2	3	4	5	6	7	
全くそう思わない	○	○	○	○	○	○	○	非常にそう思う

参考文献

[1] 松薗美帆，草野孔希 : はじめての UX リサーチ ユーザーとともに価値あるサービスを作り続けるために，翔泳社，2021.

[2] 木浦幹雄 : デザインリサーチの教科書，ビー・エヌ・エヌ新社，2020.

[3] Priede, Camilla., and Stephen, Farrall.: Comparing results from different styles of cognitive interviewing:'verbal probing'vs.'thinking aloud', International Journal of Social Research Methodology 14.4, pp. 271-287, 2011

[4] Boren, Ted, and Judith, Ramey.: Thinking aloud: Reconciling theory and practice, IEEE transactions on professional communication 43.3, pp. 261-278, 2000.

[5] Olmsted-Hawala, Erica, L. et al.: Think-aloud protocols: a comparison of three think-aloud protocols for use in testing data-dissemination web sites for usability, Proceedings of the SIGCHI conference on human factors in computing systems, pp. 2381-2390, 2010.

[6] Krahmer, Emiel, and Nicole, Ummelen.: Thinking about thinking aloud: A comparison of two verbal protocols for usability testing, IEEE transactions on professional communication 47.2, pp. 105-117, 2004.

[7] DiSalvo, Carl, and Francine, Gemperle.: From seduction to fulfillment: the use of anthropomorphic form in design, Proceedings of the 2003 international conference on Designing pleasurable products and interfaces. pp. 67-72, 2003.

[8] 大澤博隆，大村廉，今井倫太：直接擬人化手法を用いた機器からの情報提示の評価，ヒューマンインターフェース学会論文誌，Vol.10, No.3, pp.305-314, 2008.

[9] Iwamoto, Takuya, et al.: The Effectiveness of Self-Recommending Agents in Advancing Purchase Behavior Steps in Retail Marketing, Proceedings of the 9th International Conference on Human-Agent Interaction, pp. 209-217, 2021.

[10] Coefont.: https://coefont.cloud

[11] WinEyes.: https://ja.osdn.net/projects/sfnet_wineyes/

[12] muro.: https://www.deviantart.com/muro/

[13] 8bit paint.: https://minordaimyo.net/8bitpaintweb/

[14] CLOVA Note.: https://clovanote.line.me

[15] 海保博之, 原田悦子 : プロトコル分析入門―発話データから何を読むか, 新曜社, 1993.

[16] D.A.Norman.: The Psychology of Everyday Things, Basic Books, 1988.

[17] 安藤昌也 : UX デザインの教科書, 丸善出版, 2016.

[18] 蔵田洋平, 松下仁美, 小田原雄紀, 大島直樹, P. Ravindra De Silva, 岡田美智男 : Talking-Ally:聞き手性をリソースとする発話生成システムの実現にむけて, ヒューマンインターフェース学会論文誌, Vol.17, No.2, pp.159-170, 2015.