Humanization of Virtual Assistants and Delegation Choices

Nanyin Yang,[1] Marco A. Palma[1] and Andreas C. Drichoutis,[2]

[1]Texas A&M University, [2]Agricultural University of Athens

Overview

1. Introduction and Motivation
2. Research Question
3. Related Literature
4. Experimental Design
   - The Game
   - Model and Predictions
5. Study 1: Influence of VA's Gender on Delegation
6. Study 2: Interactions with VA and subsequent choice of HA
7. Summary and Conclusions

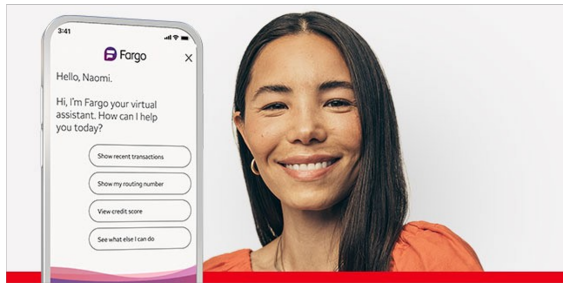
Introduction and Motivation











Introduction and Motivation

- The rise of Artificial Intelligence as omnipresent artifacts to facilitate everyday activities.

Introduction and Motivation

- The rise of Artificial Intelligence as omnipresent artifacts to facilitate everyday activities.
- Many Virtual Assistants (VA) have anthropomorphic gender features varying in realism, but they tend to be predominantly female. Why?

Introduction and Motivation

- The rise of Artificial Intelligence as omnipresent artifacts to facilitate everyday activities.
- Many Virtual Assistants (VA) have anthropomorphic gender features varying in realism, but they tend to be predominantly female. Why?
- Users may perceive a female VA as more friendly, nurturing, and empathetic, aligning (reinforcing or perpetuating) with societal stereotypes.

Research Questions

Study 1:

1. How assigning gender attributes to VA may impact individuals' willingness to delegate a decision-making search task with economic consequences?

Research Questions

Study 1:

1. How assigning gender attributes to VA may impact individuals' willingness to delegate a decision-making search task with economic consequences?
2. If there are any effects of VA gender, do they align with the patterns observed in interactions with Human Assistants?

Research Questions

Study 1:

1. How assigning gender attributes to VA may impact individuals' willingness to delegate a decision-making search task with economic consequences?
2. If there are any effects of VA gender, do they align with the patterns observed in interactions with Human Assistants?

Study 2:

3. Do interactions with gendered VA (of high or low quality) raise concerns for reinforcing prevalent gender stereotypes, potentially affecting subsequent interactions with human assistants (HA).

Related Literature

- People assign gender in synthetic high/low pitched voices (Nass et al., 1997)

Related Literature

- People assign gender in synthetic high/low pitched voices (Nass et al., 1997)
- Gender characteristics in robotic design (Nomura, 2016).

Related Literature

- People assign gender in synthetic high/low pitched voices (Nass et al., 1997)
- Gender characteristics in robotic design (Nomura, 2016).
- Strategic advice provision in repeated human-agent interactions (Azaria et al., 2012), with similar use of advice from humans and AI suggestions (Vodrahalli et al., 2021).

Related Literature

- People assign gender in synthetic high/low pitched voices (Nass et al., 1997)
- Gender characteristics in robotic design (Nomura, 2016).
- Strategic advice provision in repeated human-agent interactions (Azaria et al., 2012), with similar use of advice from humans and AI suggestions (Vodrahalli et al., 2021).
- Increase use of AI to aid in critical domains, such as medicine and finance (Agarwal et al., 2023, Baldauf et al., 2020).

Related Literature

- People assign gender in synthetic high/low pitched voices (Nass et al., 1997)
- Gender characteristics in robotic design (Nomura, 2016).
- Strategic advice provision in repeated human-agent interactions (Azaria et al., 2012), with similar use of advice from humans and AI suggestions (Vodrahalli et al., 2021).
- Increase use of AI to aid in critical domains, such as medicine and finance (Agarwal et al., 2023, Baldauf et al., 2020).
- Algorithm aversion —less tolerance to AI mistakes compared to human mistakes (Dietvorst et al., 2023).

Related Literature

- People assign gender in synthetic high/low pitched voices (Nass et al., 1997)
- Gender characteristics in robotic design (Nomura, 2016).
- Strategic advice provision in repeated human-agent interactions (Azaria et al., 2012), with similar use of advice from humans and AI suggestions (Vodrahalli et al., 2021).
- Increase use of AI to aid in critical domains, such as medicine and finance (Agarwal et al., 2023, Baldauf et al., 2020).
- Algorithm aversion —less tolerance to AI mistakes compared to human mistakes (Dietvorst et al., 2023).
- Link to a richer literature about autonomy and delegation (for example, Fehr et al., 2013; Ertac et al, 2020).

Implementation: Card Picking Game

- Authority delegation game of Fehr, Herz, and Wilkening (2013).

Implementation: Card Picking Game

- Authority delegation game of Fehr, Herz, and Wilkening (2013).
- 34 shuffled cards faced down + 1 revealed:
  - Green Card: 10
  - Blue Card: 24
  - Red Card: 40
  - Blank Cards: 0

Implementation: Card Picking Game

- Authority delegation game of Fehr, Herz, and Wilkening (2013).
- 34 shuffled cards faced down + 1 revealed:
  - Green Card: 10
  - Blue Card: 24
  - Red Card: 40
  - Blank Cards: 0
- Position of Green Card is always revealed;

Implementation: Card Picking Game

- Authority delegation game of Fehr, Herz, and Wilkening (2013).
- 34 shuffled cards faced down + 1 revealed:
  - Green Card: 10
  - Blue Card: 24
  - Red Card: 40
  - Blank Cards: 0
- Position of Green Card is always revealed;
- Successful search uncovered positions of all cards

Implementation: Card Picking Game

- Authority delegation game of Fehr, Herz, and Wilkening (2013).
- 34 shuffled cards faced down + 1 revealed:
  - Green Card: 10
  - Blue Card: 24
  - Red Card: 40
  - Blank Cards: 0
- Position of Green Card is always revealed;
- Successful search uncovered positions of all cards
- Self-search: Participants selects a card (Red Card)

Implementation: Card Picking Game

- Authority delegation game of Fehr, Herz, and Wilkening (2013).
- 34 shuffled cards faced down + 1 revealed:
  - Green Card: 10
  - Blue Card: 24
  - Red Card: 40
  - Blank Cards: 0
- Position of Green Card is always revealed;
- Successful search uncovered positions of all cards
- Self-search: Participants selects a card (Red Card)
- Assistant search: Assistant always selects Blue Card

Screenshot of Unsuccessful Search

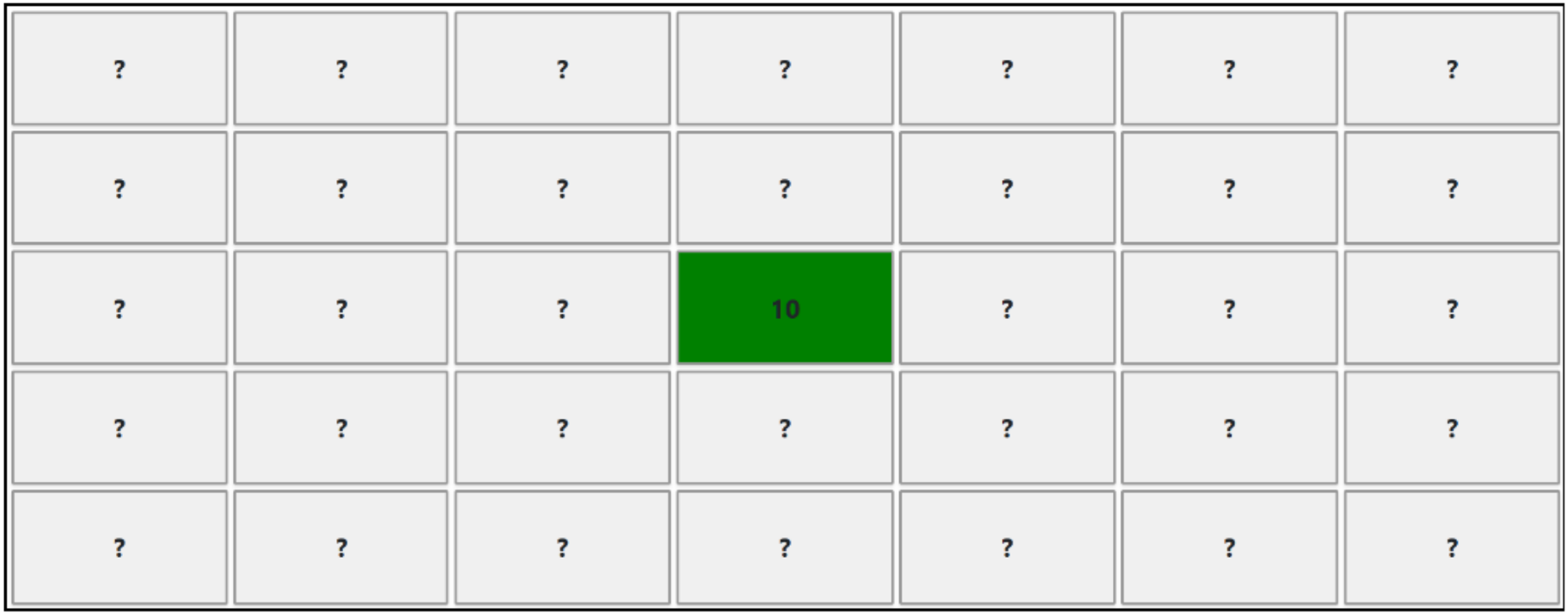


Figure: Unsuccessful Search: Only Green Card's position is revealed

Screenshot of Successful Search

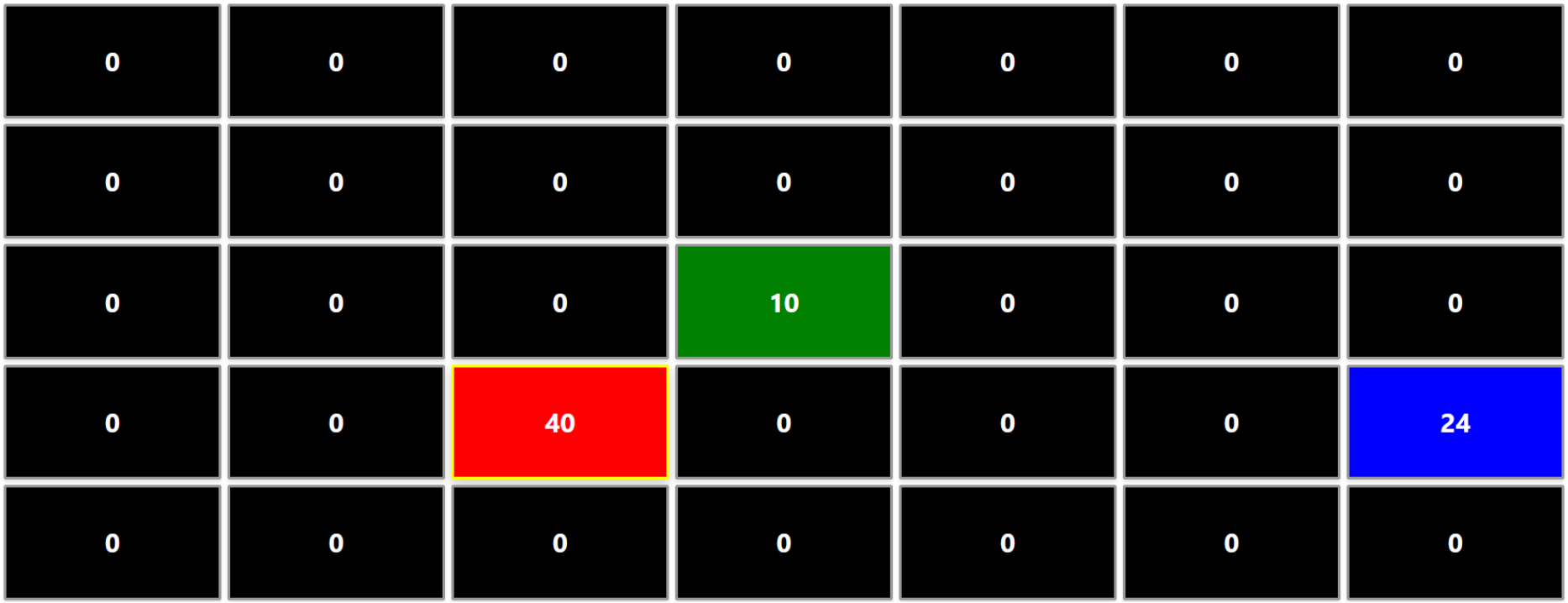


Figure: Successful Search: All cards positions are revealed

Model and Predictions

- Information search with some probability of being successful;

Model and Predictions

- Information search with some probability of being successful;
- Return from an unsuccessful search: Green Card = 10 tokens;

Model and Predictions

- Information search with some probability of being successful;
- Return from an unsuccessful search: Green Card = 10 tokens;
- Choice between two ways to conduct a search:

Model and Predictions

- Information search with some probability of being successful;
- Return from an unsuccessful search: Green Card = 10 tokens;
- Choice between two ways to conduct a search:

1. **Self-search**:
   - Choice of search intensity $E \in \{0, 0.1, 0.2, ..., 0.9, 1\}$;
   - Cost of search $= 25 \times E^2$;
   - Return from successful search: Red Card = 40 tokens.

Model and Predictions

- Information search with some probability of being successful;
- Return from an unsuccessful search: Green Card = 10 tokens;
- Choice between two ways to conduct a search:

1. **Self-search**:
   - Choice of search intensity $E \in \{0, 0.1, 0.2, ..., 0.9, 1\}$;
   - Cost of search $= 25 \times E^2$;
   - Return from successful search: Red Card = 40 tokens.
2. **Delegating to Assistant Search**:
   - Search intensity fixed at e, which equal to either **60%** or **80%**;
   - Cost of search = 0;
   - Return from successful search: Blue Card = 24 tokens;

Model and Predictions

1. **Prediction 1**: The optimal search intensity for a self-search is:

Model and Predictions

1. **Prediction 1**: The optimal search intensity for a self-search is:
$$\partial(E \times 40 + (1 - E) \times 10 - 25 \times E^2)/\partial E \rightarrow E^* = 0.6$$

Model and Predictions

1. **Prediction 1**: The optimal search intensity for a self-search is:
$$\partial(E \times 40 + (1 - E) \times 10 - 25 \times E^2)/\partial E \rightarrow E^* = 0.6$$
2. **Prediction 2**: Optimal choice is to delegate if $e \geq \frac{9}{14}$ and to self-search if $e < \frac{9}{14}$

Model and Predictions

1. **Prediction 1**: The optimal search intensity for a self-search is:
$$\partial(E \times 40 + (1 - E) \times 10 - 25 \times E^2)/\partial E \rightarrow E^* = 0.6$$
2. **Prediction 2**: Optimal choice is to delegate if $e \geq \frac{9}{14}$ and to self-search if $e < \frac{9}{14}$
   - Expected payoff from self-search is:
   $E^* \times 40 + (1 - E^*) \times 10 - 25 \times E^{*2} = 19$

Model and Predictions

1. **Prediction 1**: The optimal search intensity for a self-search is:
   $$\partial(E \times 40 + (1 - E) \times 10 - 25 \times E^2)/\partial E \rightarrow E^* = 0.6$$
2. **Prediction 2**: Optimal choice is to delegate if $e \geq \frac{9}{14}$ and to self-search if $e < \frac{9}{14}$
   - Expected payoff from self-search is:
     $E^* \times 40 + (1 - E^*) \times 10 - 25 \times E^{*2} = 19$
   - Expected payoff from the assistant search with intensity e is:
     $e \times 24 + (1 - e) \times 10 = 14e + 10$

Model and Predictions

1. **Prediction 1**: The optimal search intensity for a self-search is:
   $$\partial(E \times 40 + (1 - E) \times 10 - 25 \times E^2)/\partial E \rightarrow E^* = 0.6$$
2. **Prediction 2**: Optimal choice is to delegate if $e \geq \frac{9}{14}$ and to self-search if $e < \frac{9}{14}$
   - Expected payoff from self-search is:
     $E^* \times 40 + (1 - E^*) \times 10 - 25 \times E^{*2} = 19$
   - Expected payoff from the assistant search with intensity e is:
     $e \times 24 + (1 - e) \times 10 = 14e + 10$
   - Optimal strategy for **60%** (*low-quality* assistants): **Self-search**.

Model and Predictions

1. **Prediction 1**: The optimal search intensity for a self-search is:
   $$\partial(E \times 40 + (1 - E) \times 10 - 25 \times E^2)/\partial E \rightarrow E^* = 0.6$$
2. **Prediction 2**: Optimal choice is to delegate if $e \geq \frac{9}{14}$ and to self-search if $e < \frac{9}{14}$
   - Expected payoff from self-search is:
     $E^* \times 40 + (1 - E^*) \times 10 - 25 \times E^{*2} = 19$
   - Expected payoff from the assistant search with intensity e is:
     $e \times 24 + (1 - e) \times 10 = 14e + 10$
   - Optimal strategy for **60%** (*low-quality* assistants): **Self-search**.
   - Optimal strategy for **80%** (*high-quality* assistants): **Delegate**.

Model and Predictions

1. **Prediction 1**: The optimal search intensity for a self-search is:
   $$\partial(E \times 40 + (1 - E) \times 10 - 25 \times E^2)/\partial E \rightarrow E^* = 0.6$$
2. **Prediction 2**: Optimal choice is to delegate if $e \geq \frac{9}{14}$ and to self-search if $e < \frac{9}{14}$
   - Expected payoff from self-search is:
     $E^* \times 40 + (1 - E^*) \times 10 - 25 \times E^{*2} = 19$
   - Expected payoff from the assistant search with intensity e is:
     $e \times 24 + (1 - e) \times 10 = 14e + 10$
   - Optimal strategy for **60%** (*low-quality* assistants): **Self-search**.
   - Optimal strategy for **80%** (*high-quality* assistants): **Delegate**.
3. **Note**: The Assistant *Type* or *Gender* does not affect the optimal choices

Study 1: Influence of VA's Gender on Delegation (compared to HA)

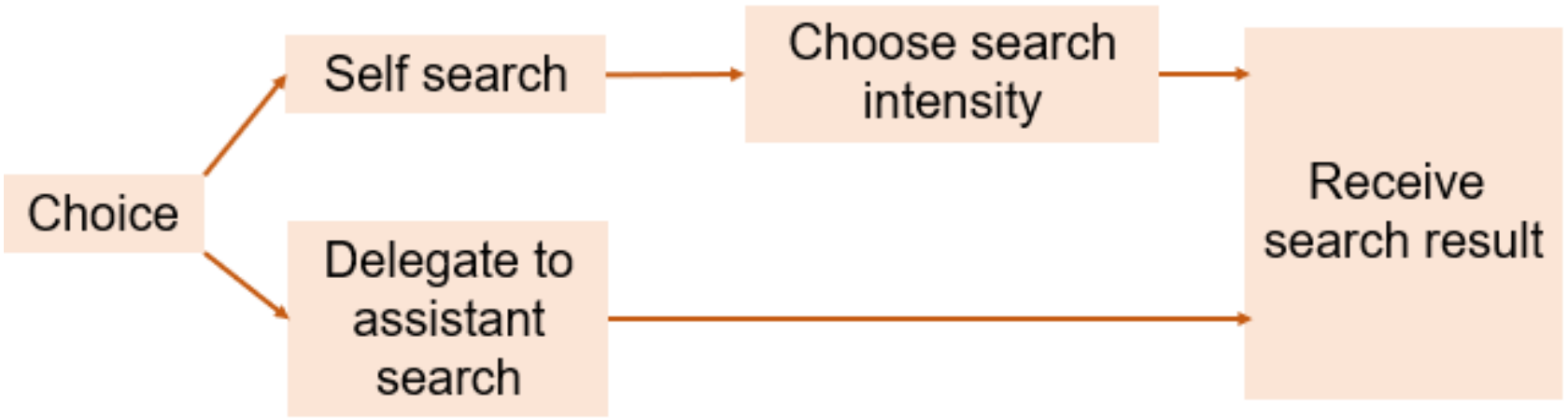


Figure: Game Procedures

Treatment Conditions

2×3 between-subject design

Treatment Conditions

2×3 between-subject design

1. Type of assistant:
   1. Virtual Assistant
   2. Human Assistant

Treatment Conditions

2×3 between-subject design

1. Type of assistant:
   1. Virtual Assistant
   2. Human Assistant
2. Gender characteristics of assistant:

Treatment Conditions

2×3 between-subject design

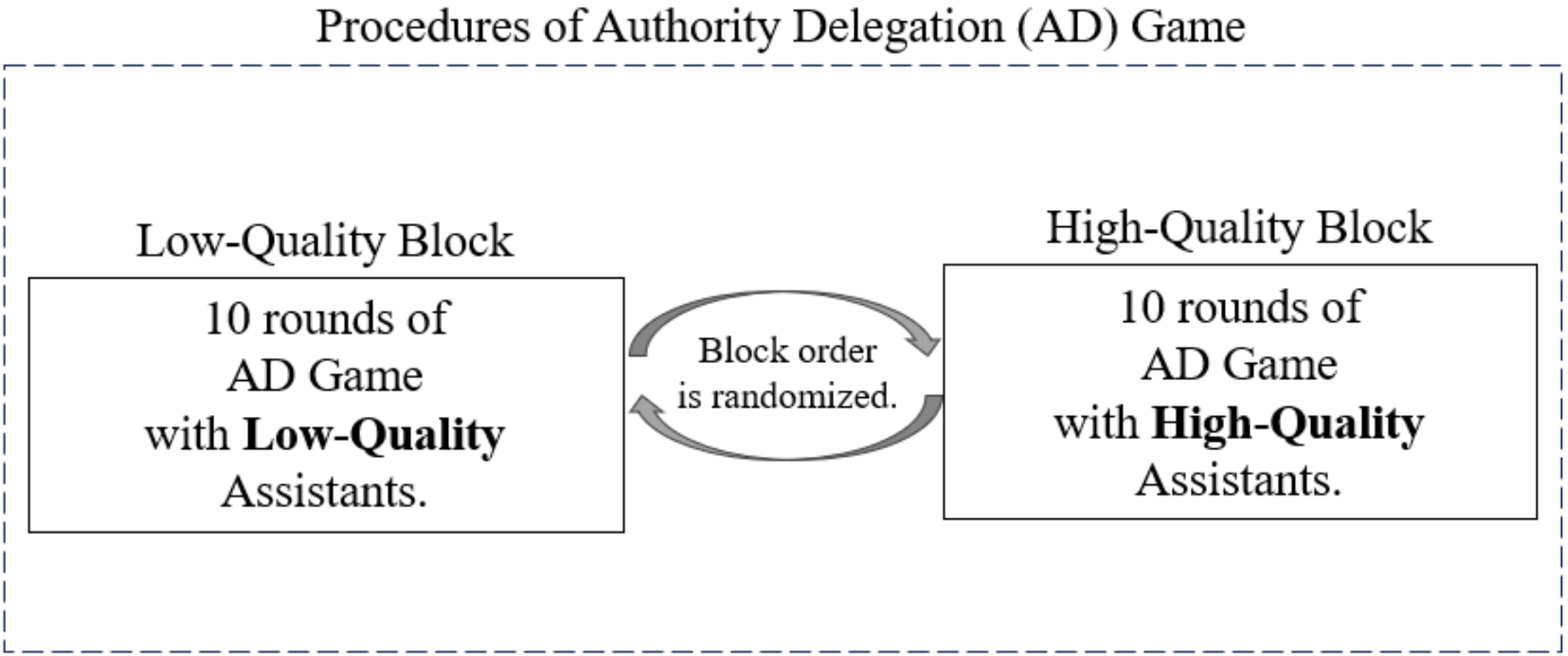
1. Type of assistant:
   1. Virtual Assistant
   2. Human Assistant
2. Gender characteristics of assistant:
   1. Female ("Jennifer")
   2. Male ("Charles")
   3. Gender-neutral ("Assistant")

▸ Assistant Construction

Assistant's Quality

- Assistant's search intensity is fixed at e:
  - *Low-Quality* Block: $e = 60\%$;
  - *High-Quality* Block: $e = 80\%$.
- Within-subject design.

Timeline of AD Game in Study 1



Between-subject: Assistants vary in types (HA; VA) and gender characteristics (male; female; gender-less).

Experimental Procedures #1

- Online experiment using Forthright Access with 672 participants.

Experimental Procedures #1

- Online experiment using Forthright Access with 672 participants.
- Procedures:
  1. Part 1: Authority-delegation game (High and Low Quality Blocks in random order);
  2. Part 2: Risk Lottery choice game;
  3. Part 3: Implicit Association Test (Gender Bias);
  4. Post-experimental questionnaire.

Experimental Procedures #1

- Online experiment using Forthright Access with 672 participants.
- Procedures:
  1. Part 1: Authority-delegation game (High and Low Quality Blocks in random order);
  2. Part 2: Risk Lottery choice game;
  3. Part 3: Implicit Association Test (Gender Bias);
  4. Post-experimental questionnaire.
- Pre-registered on AEA Registry (AEARCTR-0010921).

Experimental Procedures #2

- Subjects were warned they would be dropped if they:

Experimental Procedures #2

- Subjects were warned they would be dropped if they:
  - Spent too little time reading important instruction pages
  - Failed attention checks more than twice

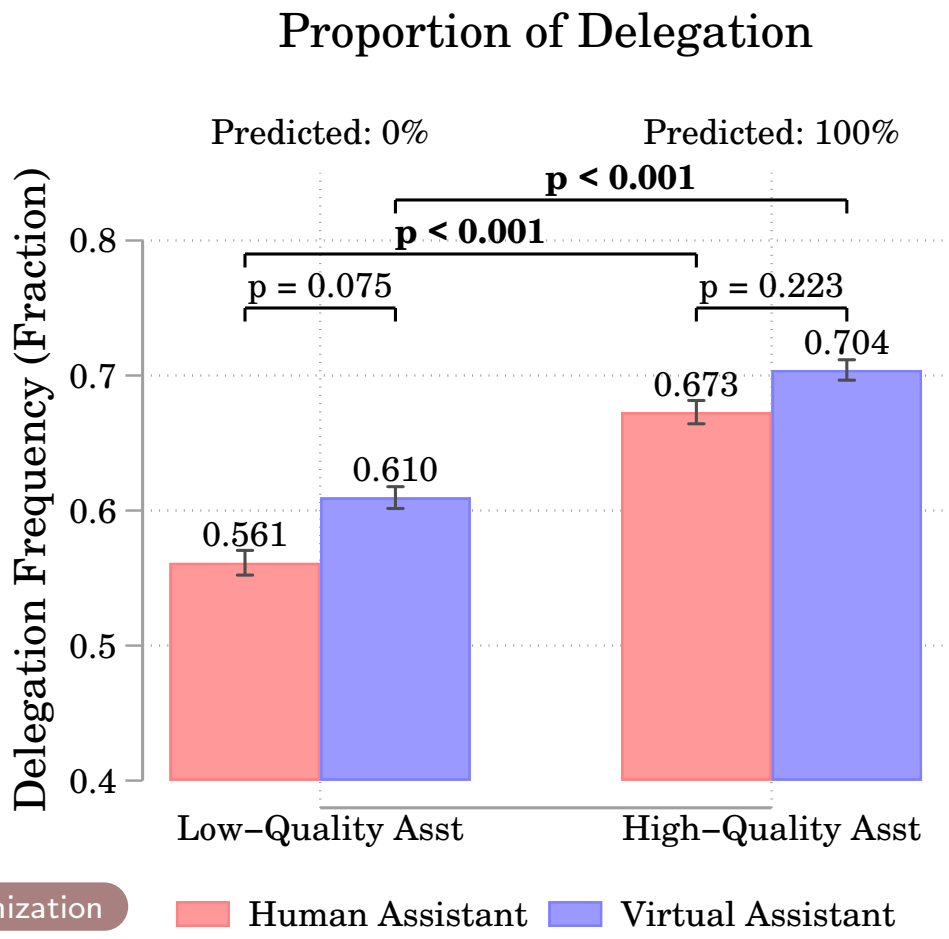
Experimental Procedures #2

- Subjects were warned they would be dropped if they:
  - Spent too little time reading important instruction pages
  - Failed attention checks more than twice
- 49% subjects qualified with total of 672

▸ Warning Page

	Treatment Conditions					
	VA Assistants			Human Assistants		
	Female	Male	Gender-less	Female	Male	Gender-less
Female Subjects (%)	45.9	48.7	50.9	48.2	47.3	51.1
Age	45.2	43.3	47.7	45.2	44.8	44.7
	(17.5)	(16.5)	(17.0)	(17.1)	(16.0)	(15.2)
N	134	115	116	110	93	92

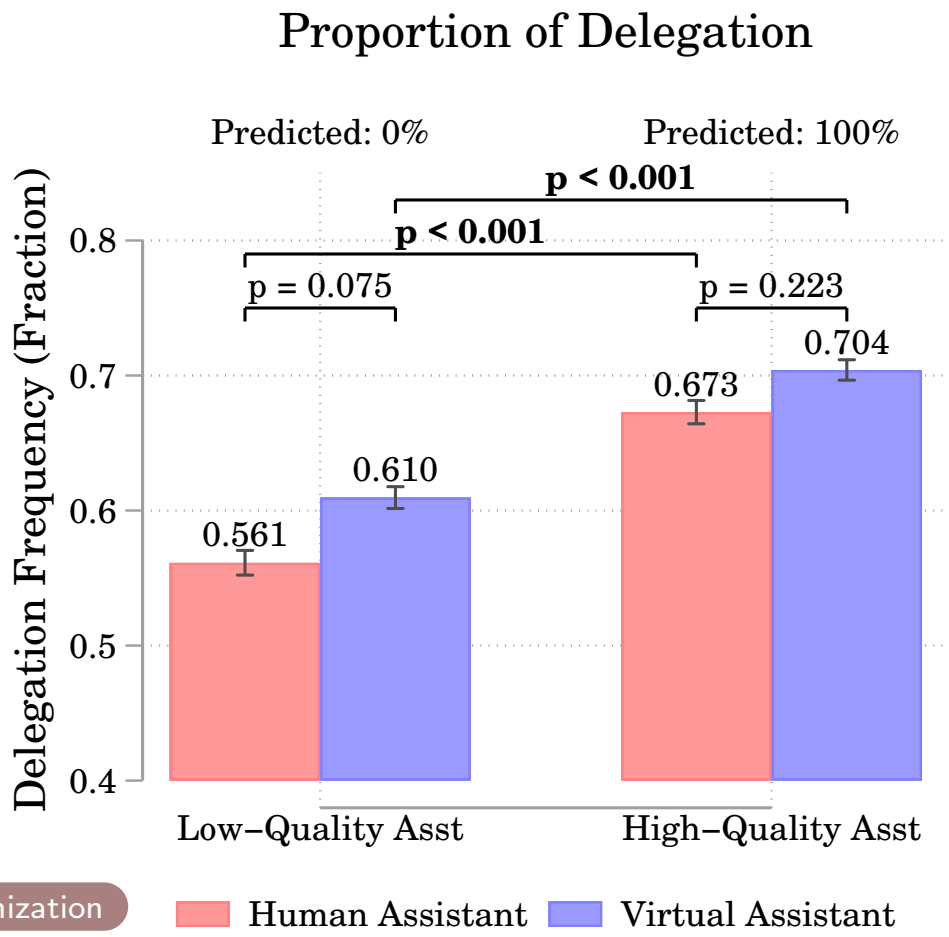
Results: Search Delegation



▸ Manipulation Check: Humanization

Results: Search Delegation

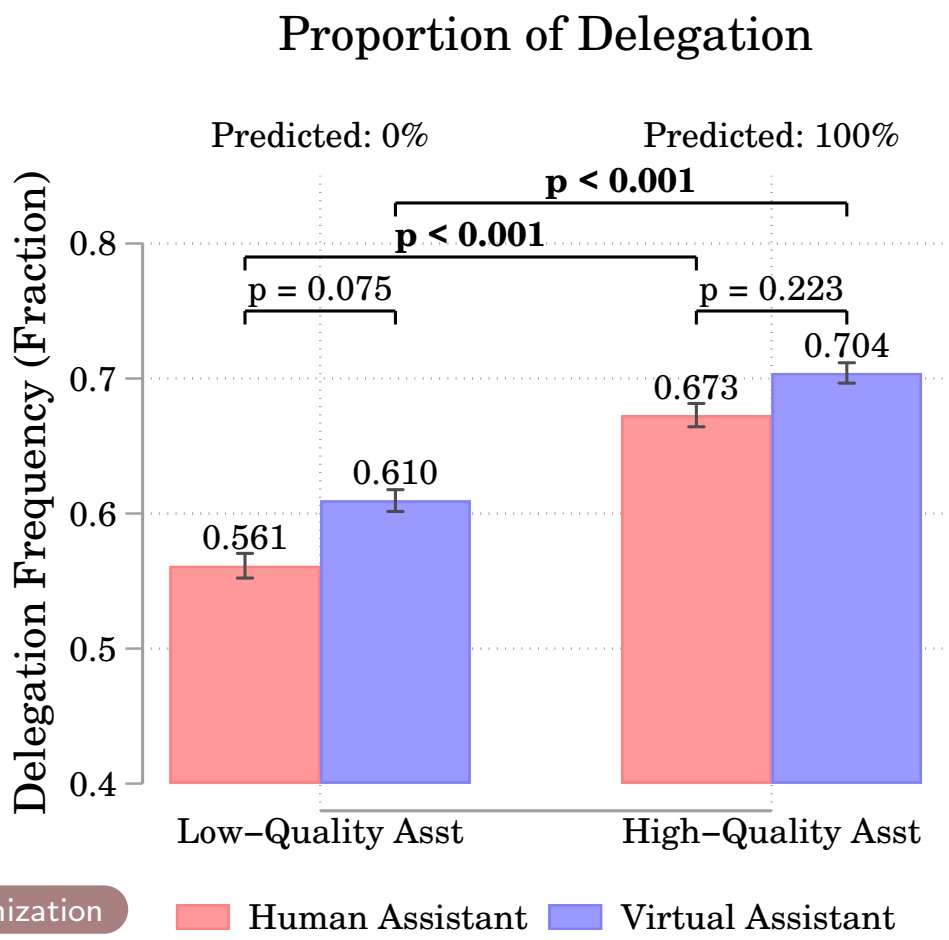
- Overdelegation for LQ; underdelegation for HQ.



▸ Manipulation Check: Humanization

Results: Search Delegation

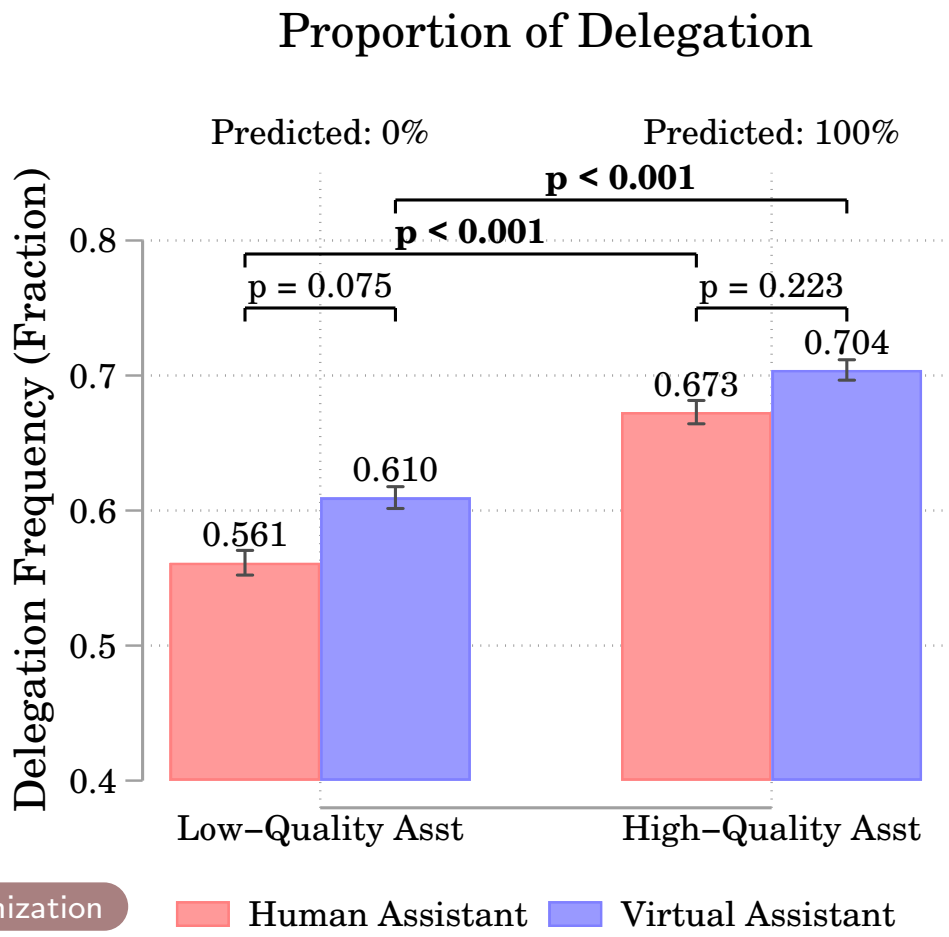
- Overdelegation for LQ; underdelegation for HQ.
- More delegation to virtual compared to human.



▸ Manipulation Check: Humanization

Results: Search Delegation

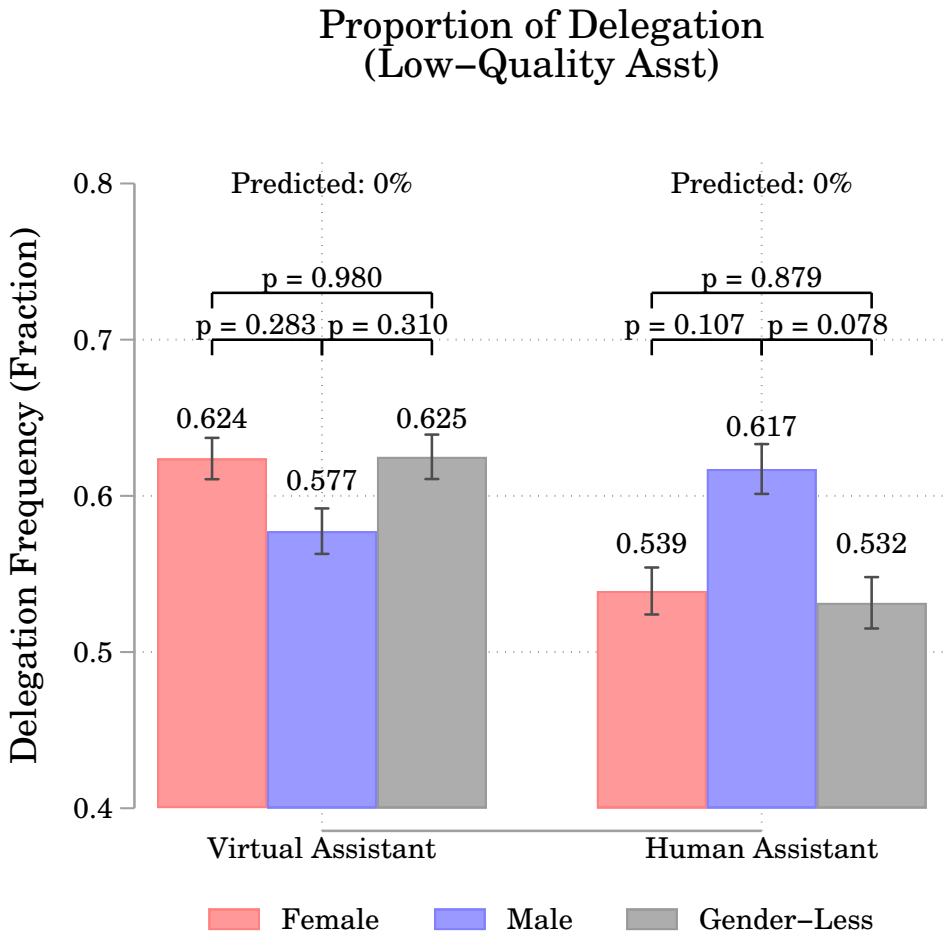
- Overdelegation for LQ; underdelegation for HQ.
- More delegation to virtual compared to human.
- More delegation to high-quality than to low-quality.



▸ Manipulation Check: Humanization

Results: Delegation to Low-Quality

- With low-quality assistants: Male is least favored among VAs; but Male is most favored among HAs.

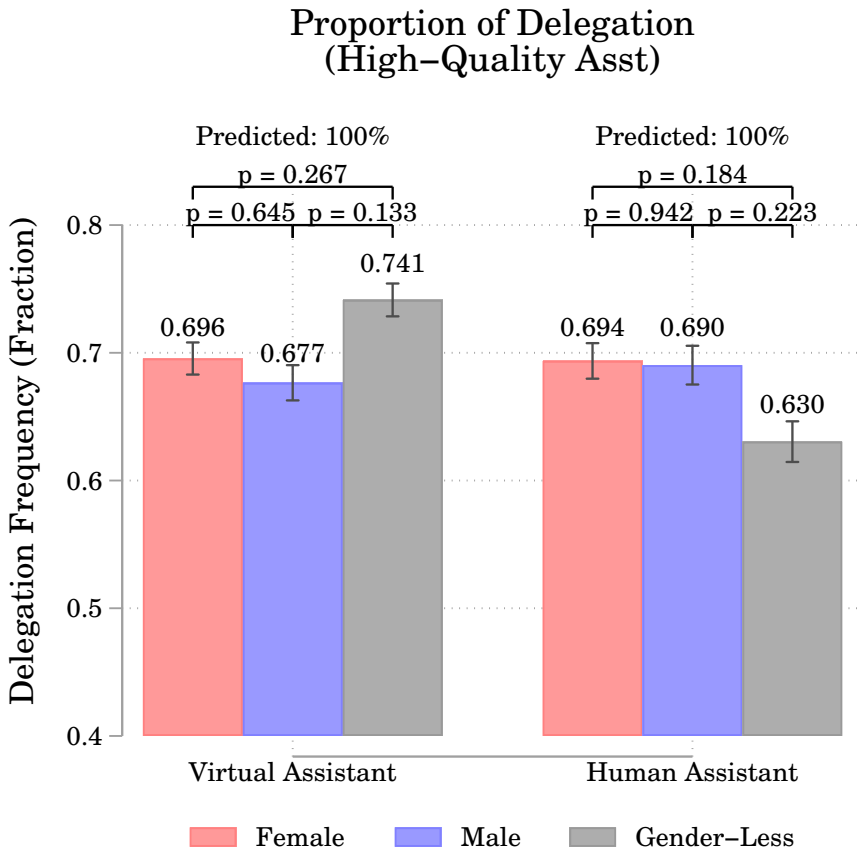


▸ Manipulation Check: VA

▸ Manipulation Check: HA

Results: Delegation to High-Quality

- Suppressing gender increases delegation for virtual assistants.



Panel Logit Reg.: Delegation to Asst

- Dependent variable:
  - Binary indicator of delegation to assistant;

Panel Logit Reg.: Delegation to Asst

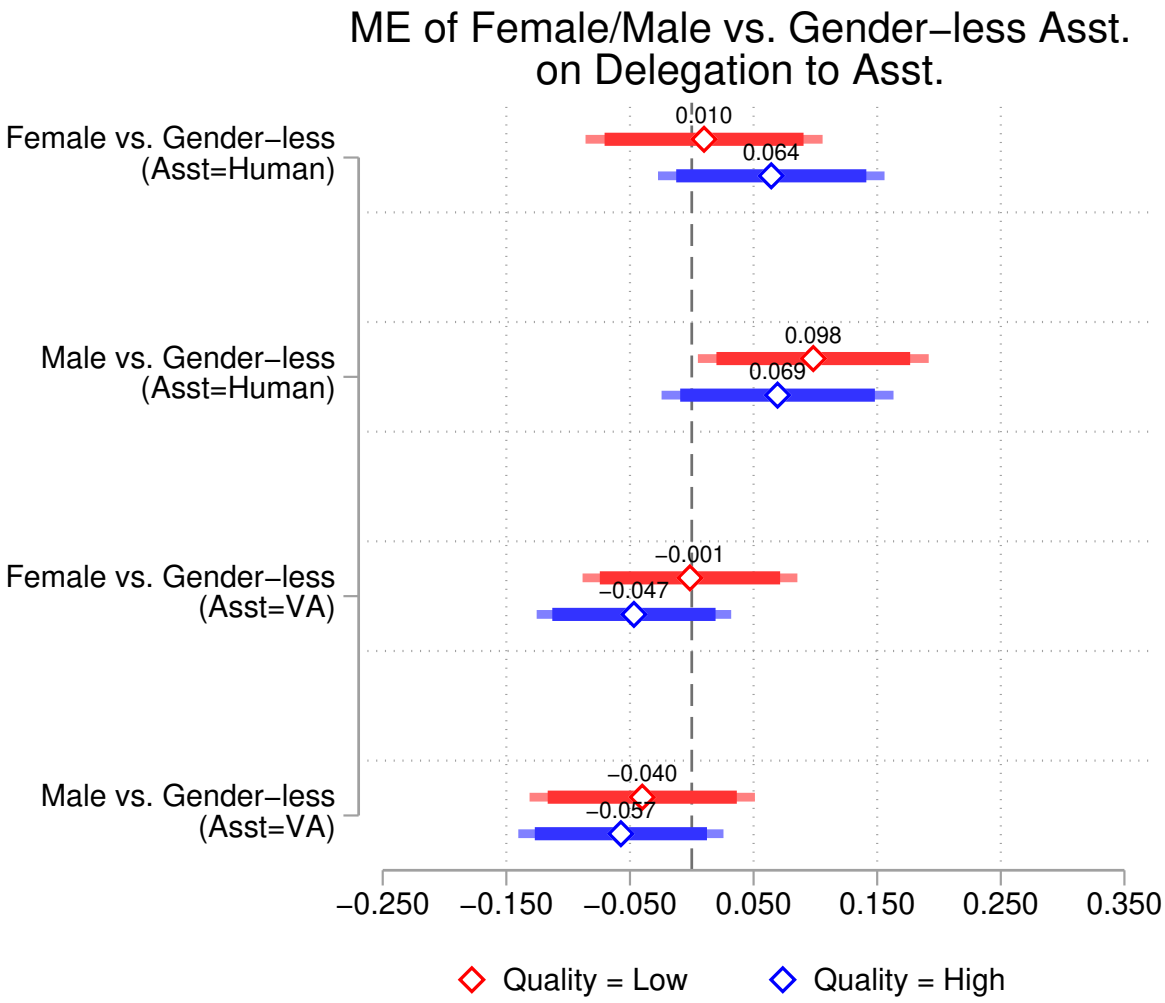
- Dependent variable:
  - Binary indicator of delegation to assistant;
- Independent variables:
  - Binary indicator of assistant as VA;
  - Indicator of assistant's gender;
  - Binary indicator of assistant as high-quality;

Panel Logit Reg.: Delegation to Asst

- Dependent variable:
  - Binary indicator of delegation to assistant;
- Independent variables:
  - Binary indicator of assistant as VA;
  - Indicator of assistant's gender;
  - Binary indicator of assistant as high-quality;
- Other controls:
  - Round; Gender; Age; Ethnicity; Risk tolerance; Trust.

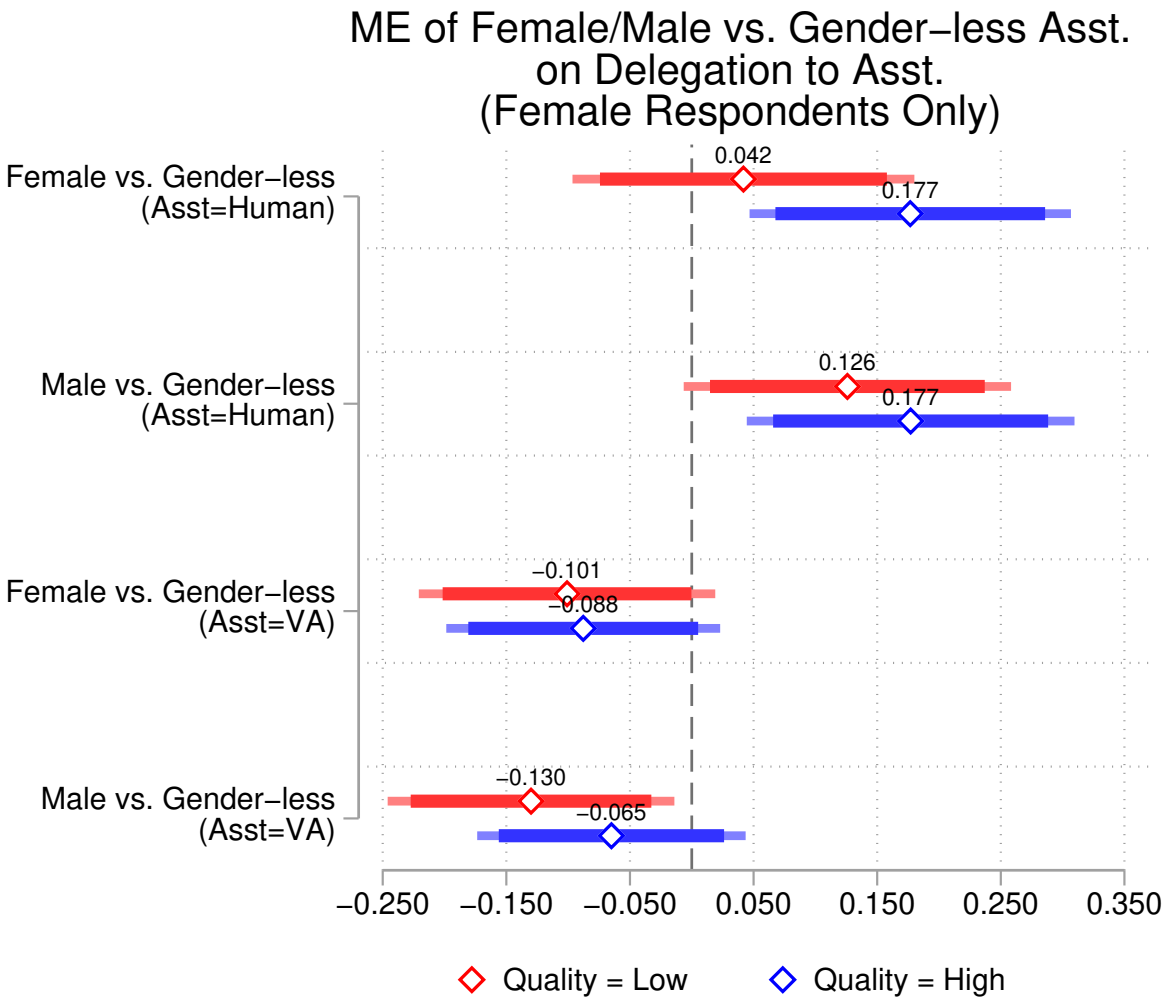
Marginal Effect of Gender on Delegation

- All respondents:



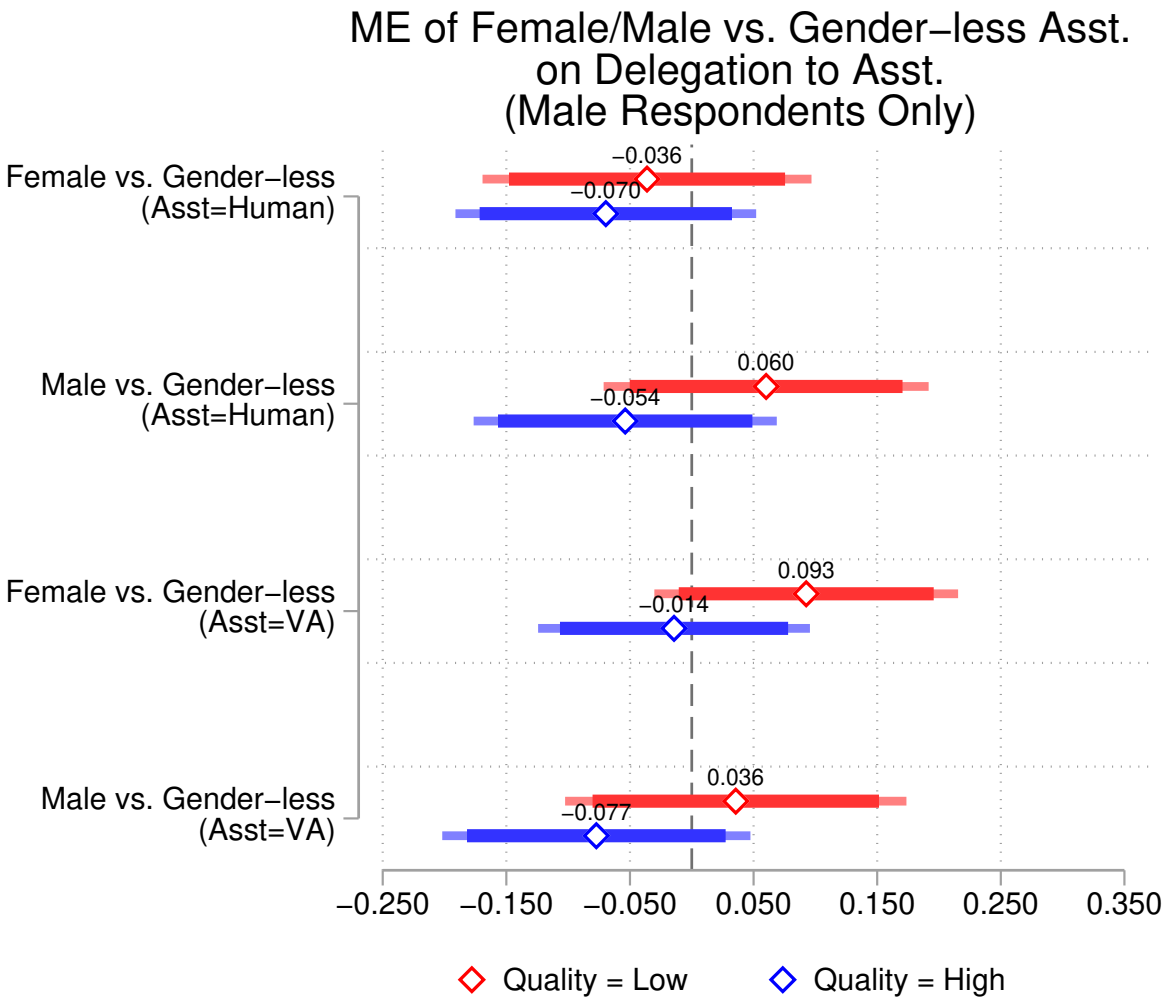
ME (Female Respondents Only)

- Female respondents reduce delegation when VA is humanized.



ME (Male Respondents Only)

- Male respondents are not responsive to assistant's gender.



Does VA Quality Shape Gender Stereotype?

- Brief IAT measures implicit association of male/female with pleasant words. ▸ BIAT
  - D-score indicator = 1: association between **female** and pleasant words;
  - D-score indicator = 0: association between **male** and pleasant words;

Does VA Quality Shape Gender Stereotype?

- Brief IAT measures implicit association of male/female with pleasant words. ▸ BIAT
  - D-score indicator = 1: association between **female** and pleasant words;
  - D-score indicator = 0: association between **male** and pleasant words;
- "High-Quality First" indicator = 1 if respondent interacts with high-quality assistants first, then low-quality.

Logit Reg.: ME of High-Quality First on IAT Indicator

	DV: Binary Indicator of Positive D-Score from Brief Implicit Association Test					
	Virtual Assistant			Human Assistant		
	(1) Female	(2) Male	(3) Gender-less	(4) Female	(5) Male	(6) Gender-less
High-Quality First	0.462***	-0.469***	0.421***	0.447***	-0.457***	0.465***
	(0.016)	(0.059)	(0.045)	(0.048)	(0.027)	(0.024)
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
N of Individuals	127	112	114	108	89	90

Logit Reg.: ME of High-Quality First on IAT Indicator

- "First impression" matters: assistant quality in the first block spillovers to gender bias captured by IAT.

	DV: Binary Indicator of Positive D-Score from Brief Implicit Association Test					
	Virtual Assistant			Human Assistant		
	(1) Female	(2) Male	(3) Gender-less	(4) Female	(5) Male	(6) Gender-less
High-Quality First	0.462***	-0.469***	0.421***	0.447***	-0.457***	0.465***
	(0.016)	(0.059)	(0.045)	(0.048)	(0.027)	(0.024)
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
N of Individuals	127	112	114	108	89	90

Study 2: Do interactions with gendered VA affect subsequent choice of HA?

- Identical experimental procedures as Study 1; N = 637.

Study 2: Do interactions with gendered VA affect subsequent choice of HA?

- Identical experimental procedures as Study 1; N = 637.
- Two blocks of AD games
  - Block 1: Interaction with **VA**
  - Block 2: Interaction with **HA**

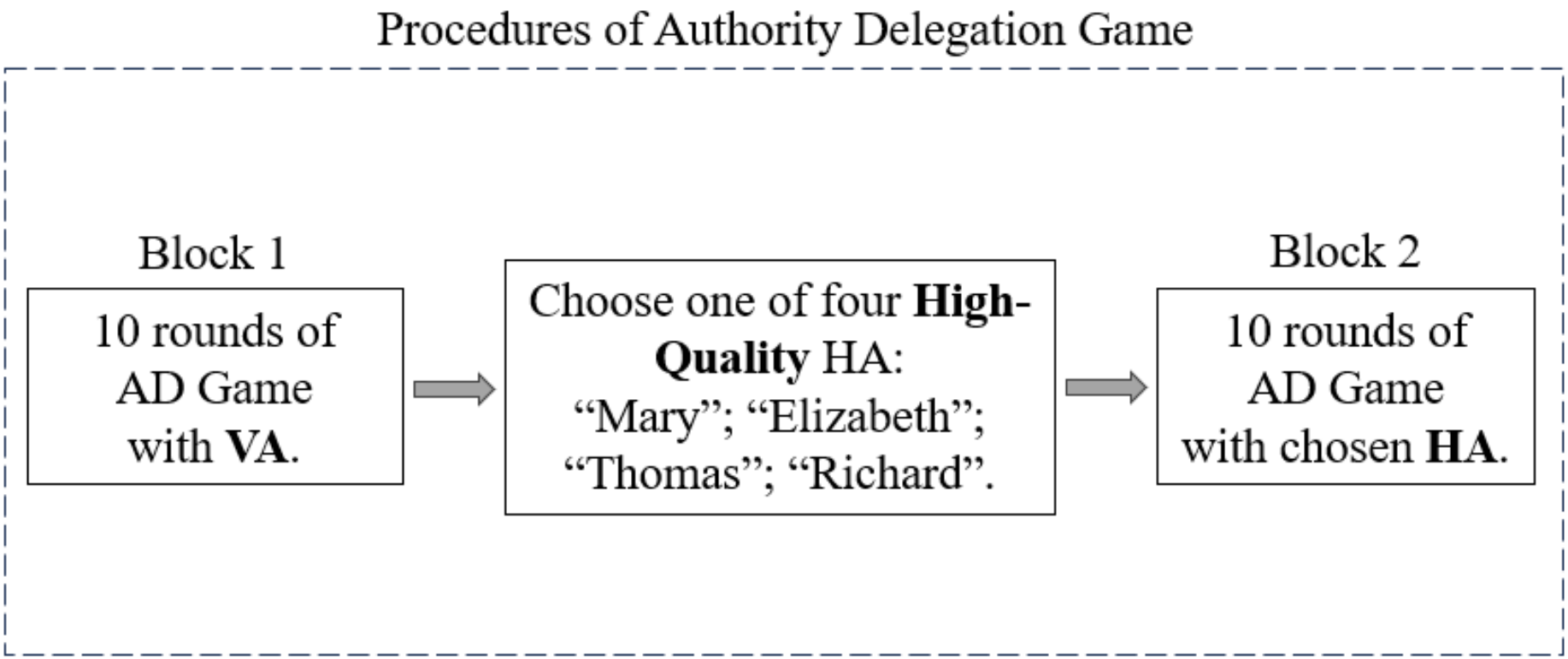
Study 2: Do interactions with gendered VA affect subsequent choice of HA?

- Identical experimental procedures as Study 1; N = 637.
- Two blocks of AD games
  - Block 1: Interaction with **VA**
  - Block 2: Interaction with **HA**
- 3×2 design of VA varying:
  - **Gender characteristics**: Male; Female; Gender-less;
  - **Quality**: High; Low.

Study 2: Do interactions with gendered VA affect subsequent choice of HA?

- Identical experimental procedures as Study 1; N = 637.
- Two blocks of AD games
  - Block 1: Interaction with **VA**
  - Block 2: Interaction with **HA**
- 3×2 design of VA varying:
  - **Gender characteristics**: Male; Female; Gender-less;
  - **Quality**: High; Low.
- Subjects **choose an HA** to interact with *before* entering Block 2.

Timeline of AD Game in Study 2



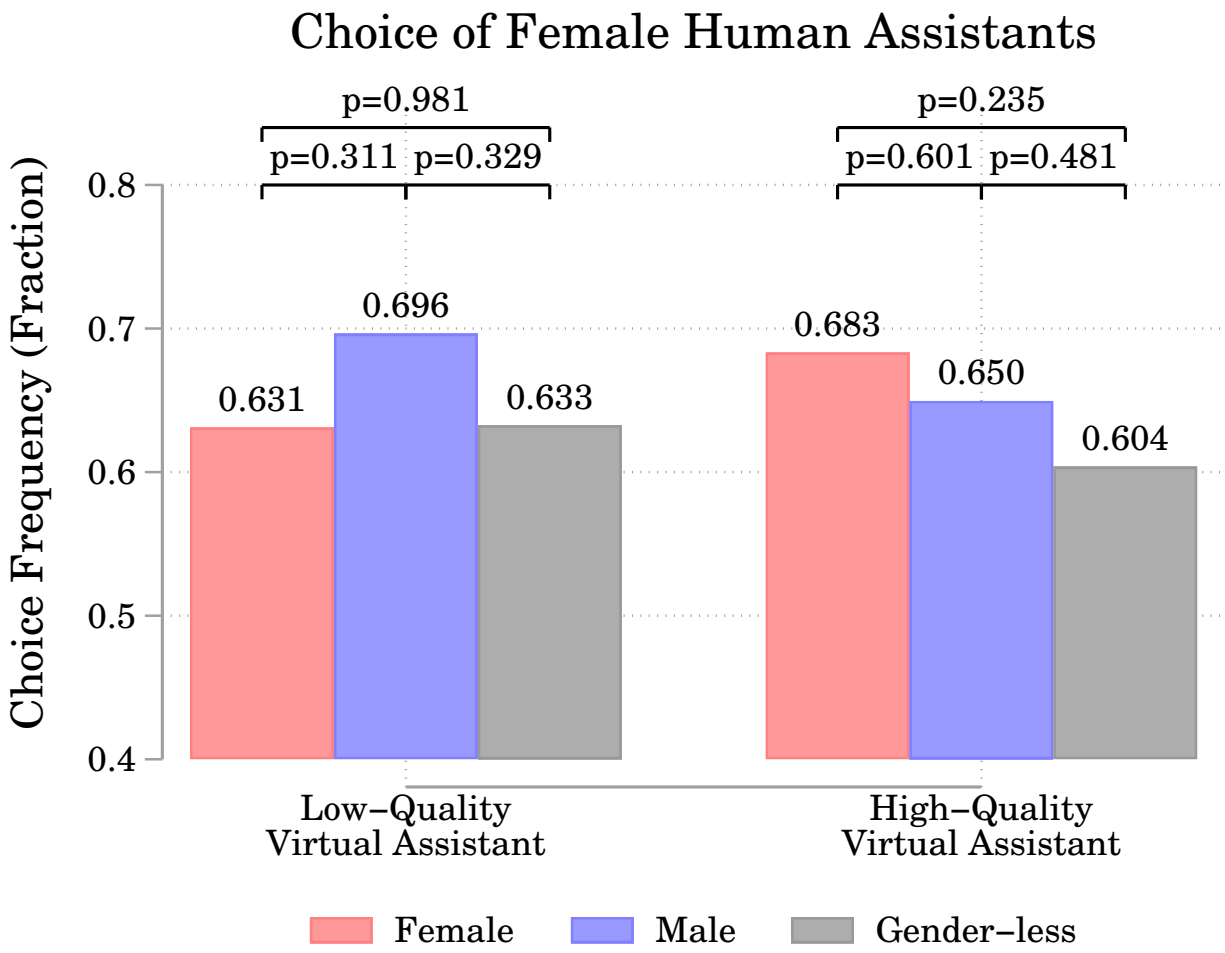
Between-subject: VA in Block 1 varies in **quality (high; low)** and **gender characteristics (male; female; gender-less)**.

Perception of VA Helpfulness

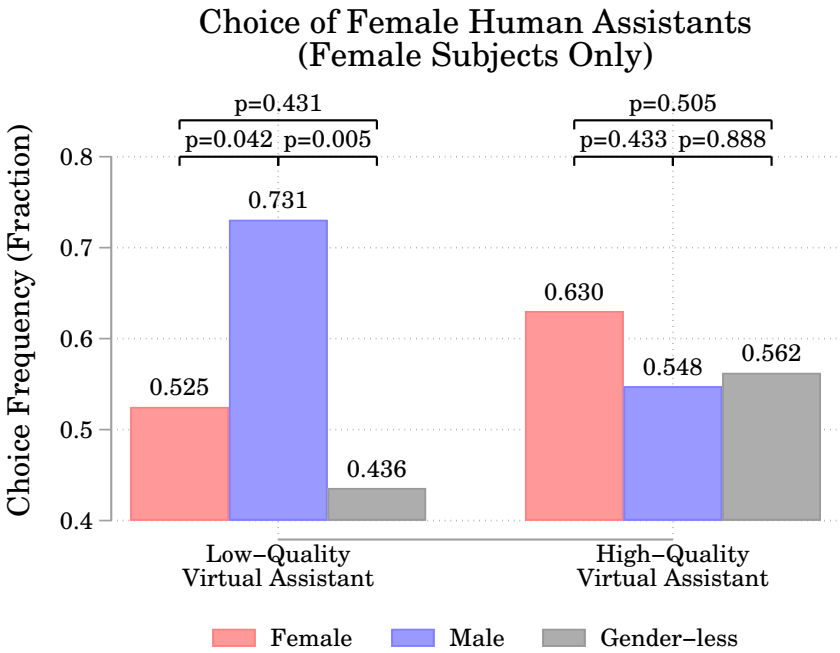
- High-quality VAs are perceived as more competent and helpful, without biases based on gender attributes.

	DV: Perceived VA Competence/Helpfulness					
	Male & Female Subjects		Female Subjects		Male Subjects	
	(1)	(2)	(3)	(4)	(5)	(6)
Female VA	0.234	0.235	0.476	0.503	0.079	0.103
	(0.227)	(0.229)	(0.342)	(0.330)	(0.302)	(0.305)
Male VA	0.253	0.246	0.321	0.289	0.200	0.202
	(0.210)	(0.210)	(0.341)	(0.343)	(0.268)	(0.267)
High-Quality VA	0.737***	0.792***	0.614*	0.619*	0.845***	0.948***
	(0.209)	(0.213)	(0.344)	(0.346)	(0.261)	(0.266)
Female VA × High-Quality VA	-0.070	-0.090	-0.213	-0.134	0.003	-0.049
	(0.287)	(0.288)	(0.438)	(0.418)	(0.378)	(0.391)
Male VA × High-Quality VA	-0.192	-0.223	-0.264	-0.196	-0.178	-0.215
	(0.281)	(0.280)	(0.461)	(0.463)	(0.354)	(0.352)
Constant	4.980***	4.959***	4.949***	4.077***	5.000***	5.435***
	(0.163)	(0.351)	(0.272)	(0.521)	(0.204)	(0.448)
Individual Controls	No	Yes	No	Yes	No	Yes
N of Individuals	637	633	267	266	370	367

Spillover from VA to Choice of HA

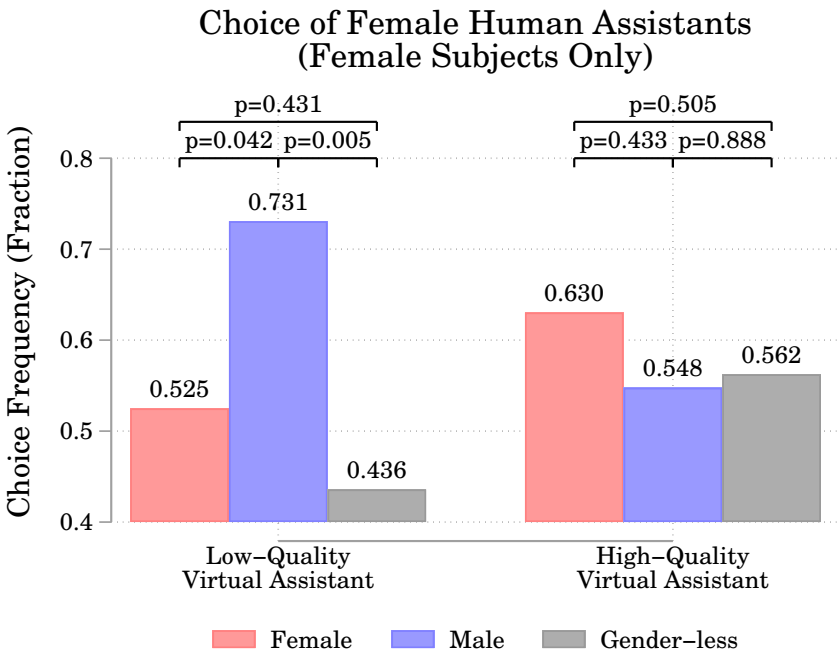


Choice of HA by Respondent Gender

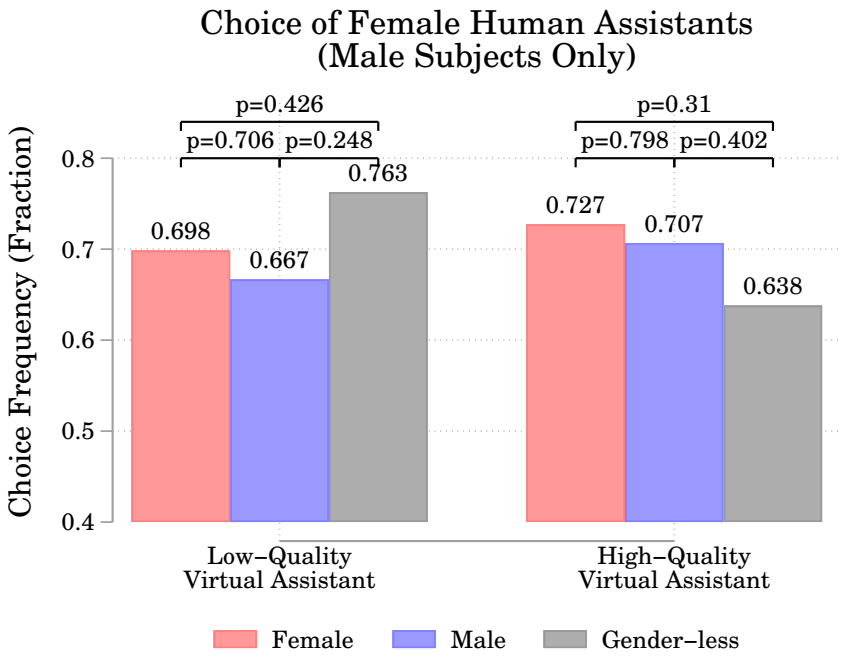


(a) Female Respondents

Choice of HA by Respondent Gender



(a) Female Respondents

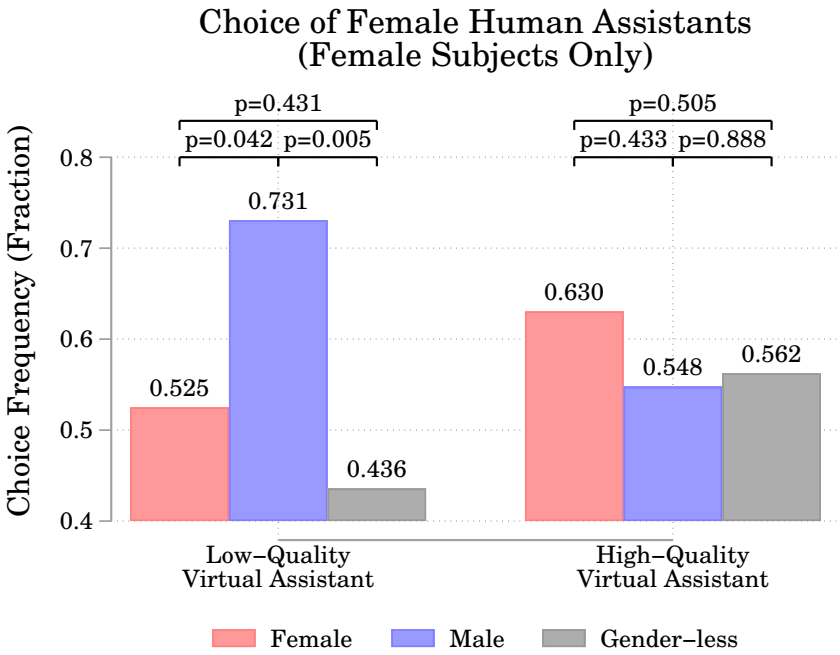


(b) Male Respondents

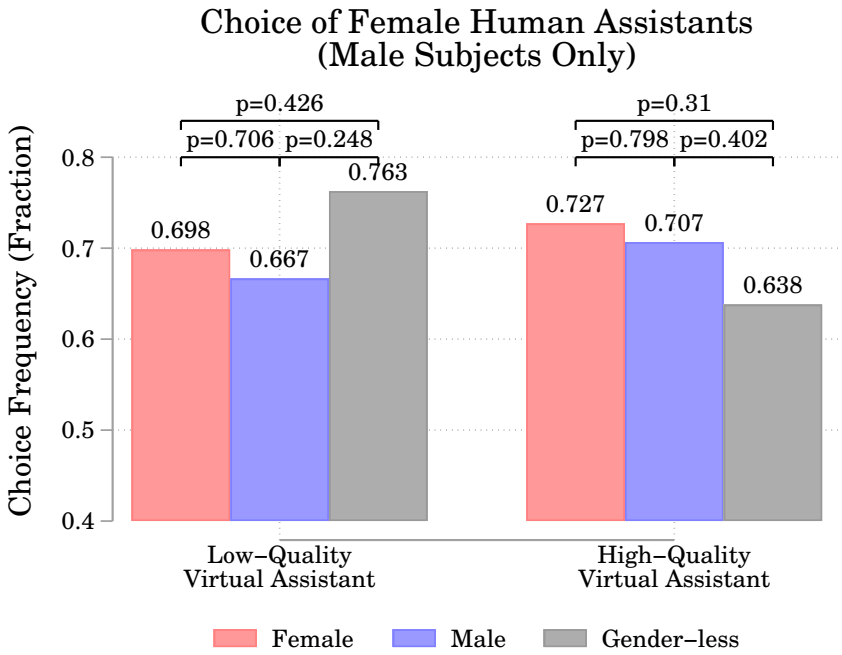
▸ ME

Choice of HA by Respondent Gender

- Female respondents are more likely to choose *female* HA after interacting with a low-quality *male* VA.



(a) Female Respondents



(b) Male Respondents

▸ ME

Summary and Concluding Remarks

1. Authority-delegation game, with three dimensions of varying conditions regarding the assistant:
   - Search intensity (i.e., low and high quality);
   - Gender characteristics (i.e., male, female, no gender; Human or AI.

Summary and Concluding Remarks

1. Authority-delegation game, with three dimensions of varying conditions regarding the assistant:
   - Search intensity (i.e., low and high quality);
   - Gender characteristics (i.e., male, female, no gender; Human or AI.
2. Key findings:
   - Low-quality assistant: Female and Gender-Neutral AI's are most delegated; But Male Humans are most delegated.

Summary and Concluding Remarks

1. Authority-delegation game, with three dimensions of varying conditions regarding the assistant:
   - Search intensity (i.e., low and high quality);
   - Gender characteristics (i.e., male, female, no gender; Human or AI.
2. Key findings:
   - Low-quality assistant: Female and Gender-Neutral AI's are most delegated; But Male Humans are most delegated.
   - High-quality assistant: Gender/Assistant's types have a smaller impact on delegation decisions.
   - Gender-related implicit association predicts delegation to both Human and AI.

Summary and Concluding Remarks

1. Authority-delegation game, with three dimensions of varying conditions regarding the assistant:
   - Search intensity (i.e., low and high quality);
   - Gender characteristics (i.e., male, female, no gender; Human or AI.
2. Key findings:
   - Low-quality assistant: Female and Gender-Neutral AI's are most delegated; But Male Humans are most delegated.
   - High-quality assistant: Gender/Assistant's types have a smaller impact on delegation decisions.
   - Gender-related implicit association predicts delegation to both Human and AI.
   - Only minor spillover effects with lower male selections after interacting with low-quality male AI.

Summary and Concluding Remarks

1. Authority-delegation game, with three dimensions of varying conditions regarding the assistant:
   - Search intensity (i.e., low and high quality);
   - Gender characteristics (i.e., male, female, no gender; Human or AI.
2. Key findings:
   - Low-quality assistant: Female and Gender-Neutral AI's are most delegated; But Male Humans are most delegated.
   - High-quality assistant: Gender/Assistant's types have a smaller impact on delegation decisions.
   - Gender-related implicit association predicts delegation to both Human and AI.
   - Only minor spillover effects with lower male selections after interacting with low-quality male AI.
   - The results of both studies seem to be driven by female respondents.

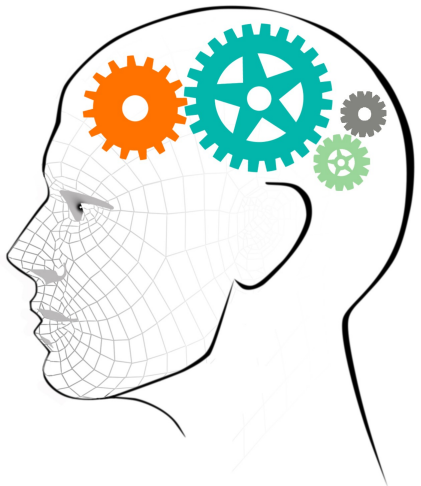
Thank you!

Do you have a paper we need to cite? Please send it to
mapalma@tamu.edu



AUA
AGRICULTURAL UNIVERSITY OF ATHENS





HUMAN BEHAVIOR LABORATORY
THE TEXAS A&M UNIVERSITY SYSTEM

Experimental Procedures – more details

- Subjects were warned they would be dropped if they:
  - Spent too little time reading important instruction pages
  - Failed attention checks more than twice
- 49% subjects qualified with total of 672

▸ Warning Page

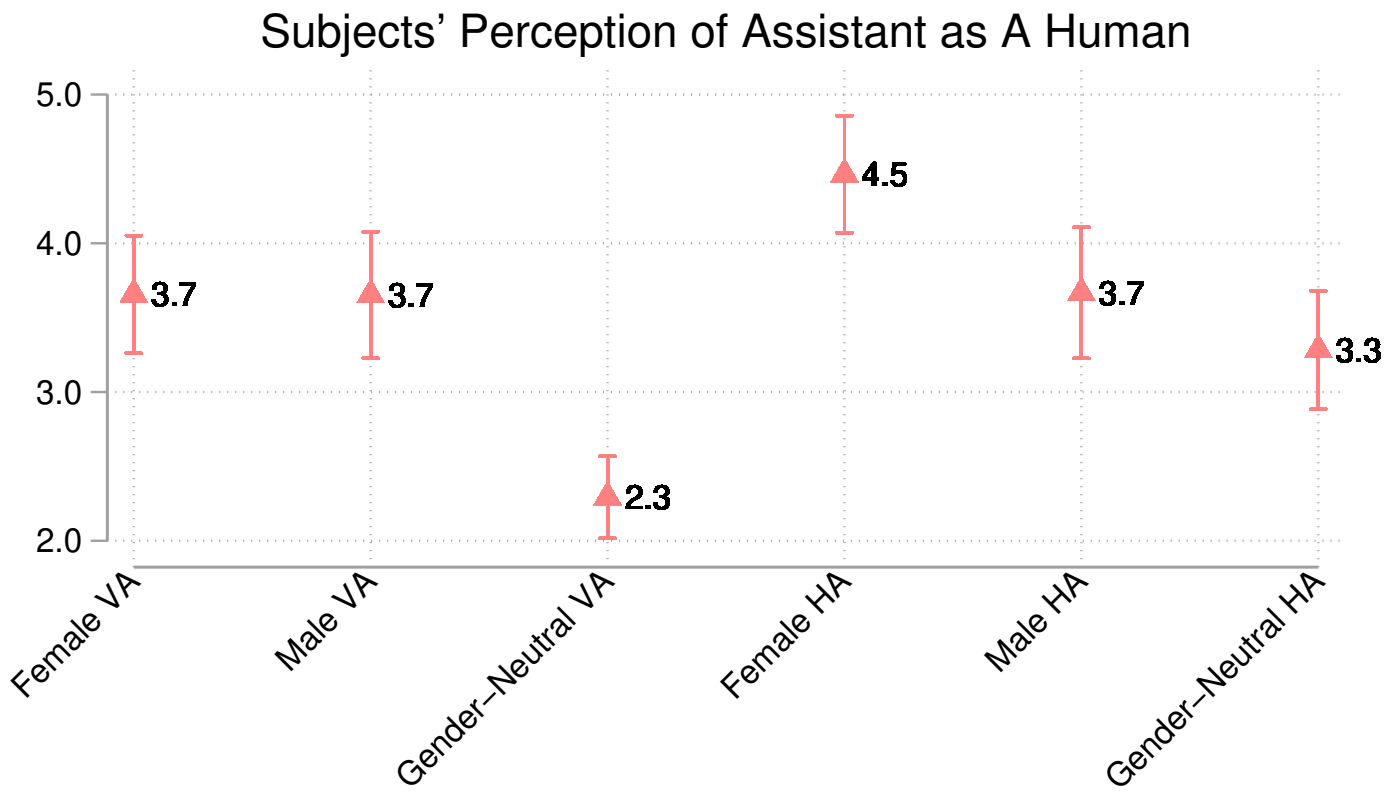
	Treatment Conditions					
	VA Assistants			Human Assistants		
	Female	Male	Gender-less	Female	Male	Gender-less
Female Subjects (%)	45.9	48.7	50.9	48.2	47.3	51.1
Age	45.2	43.3	47.7	45.2	44.8	44.7
	(17.5)	(16.5)	(17.0)	(17.1)	(16.0)	(15.2)
N	134	115	116	110	93	92

Manipulation Check

- Post-experimental survey: collect subjects' perception of the assistant as **human**, **female**, and **male**.

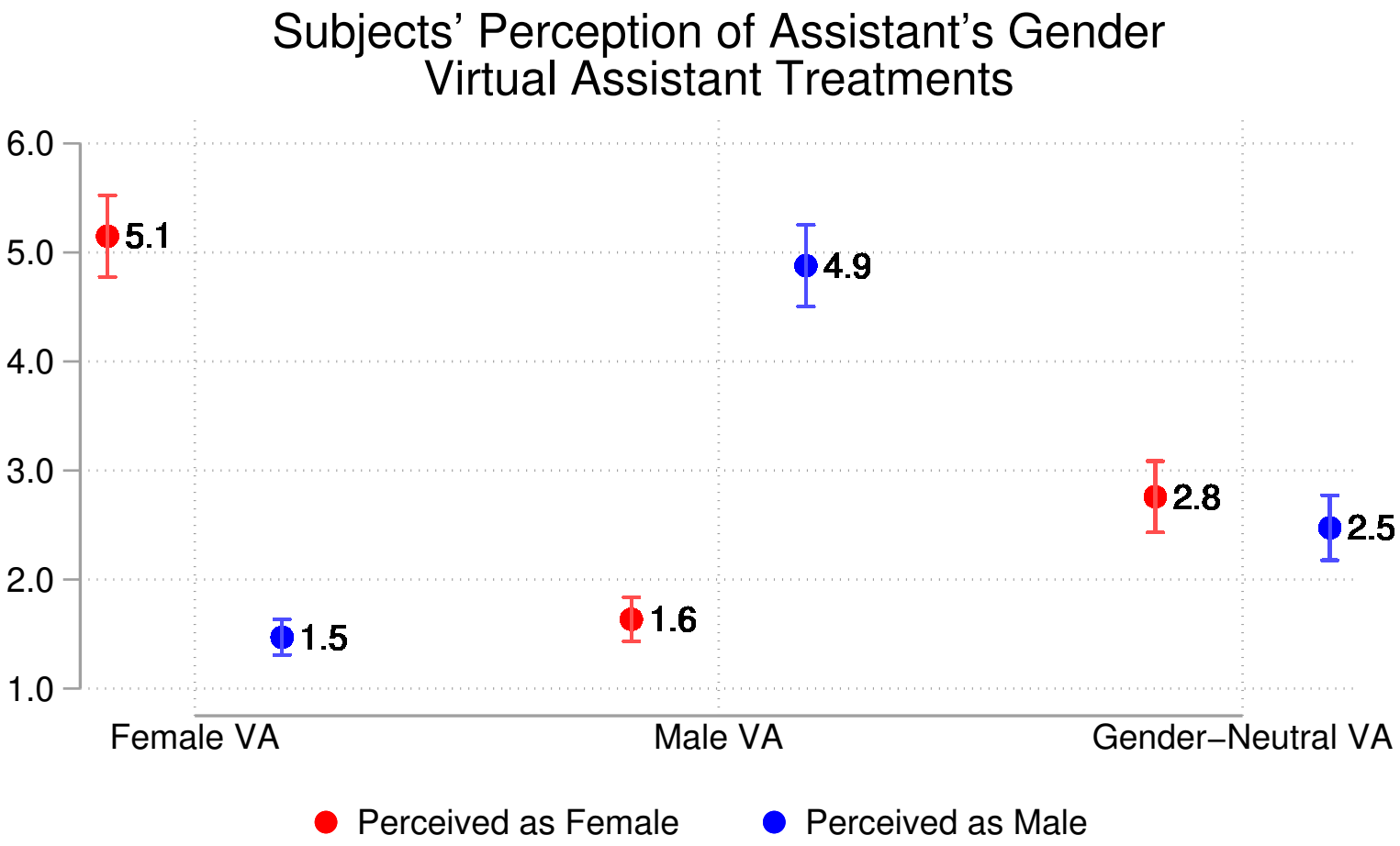
Manipulation Check: Humanization

- Giving the VA a male/female name increases subjects' perception of the VA as a human;
- Perception of HA as a human is stronger than of VA.



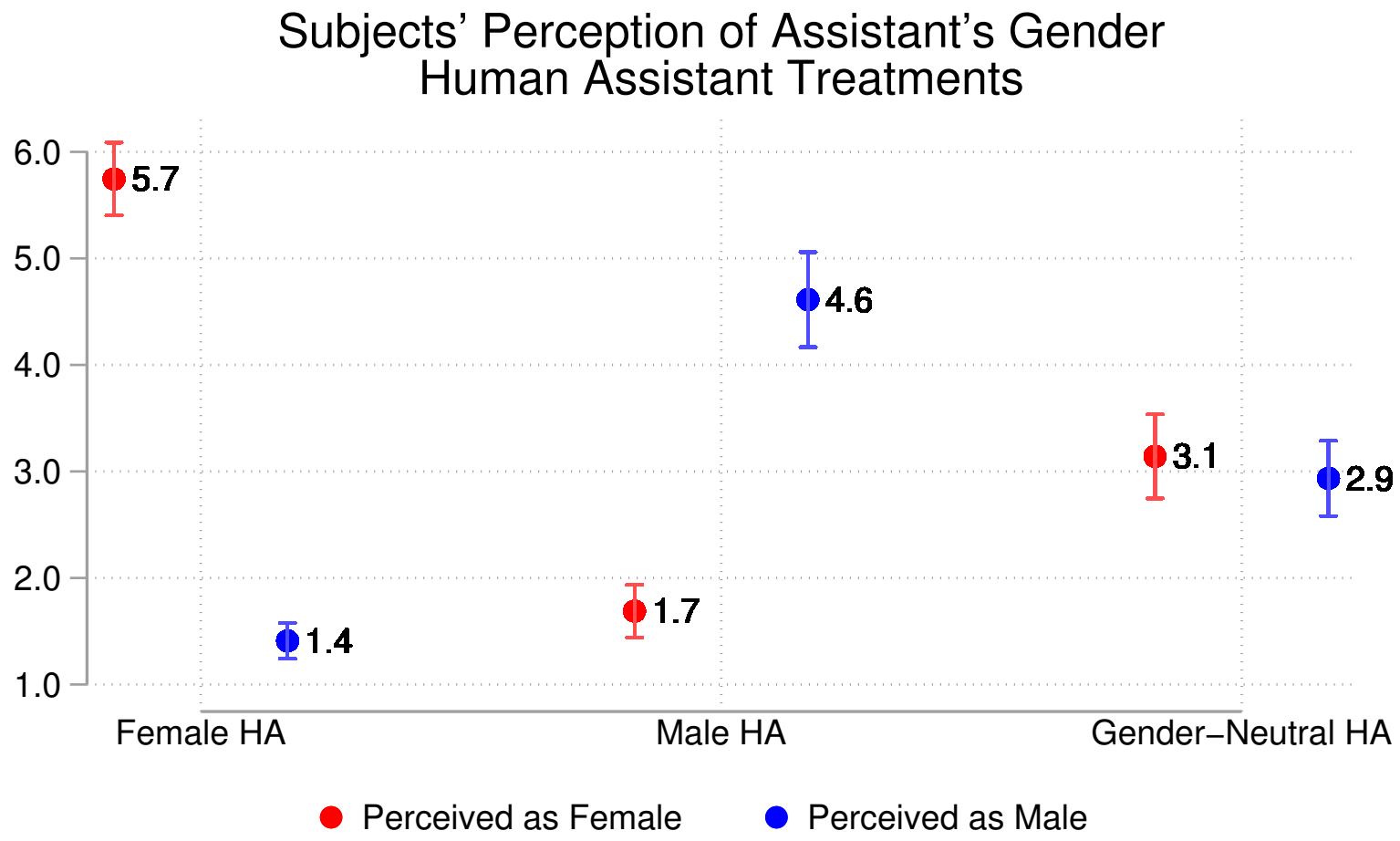
Return

Manipulation Check: Virtual Assistant



Return

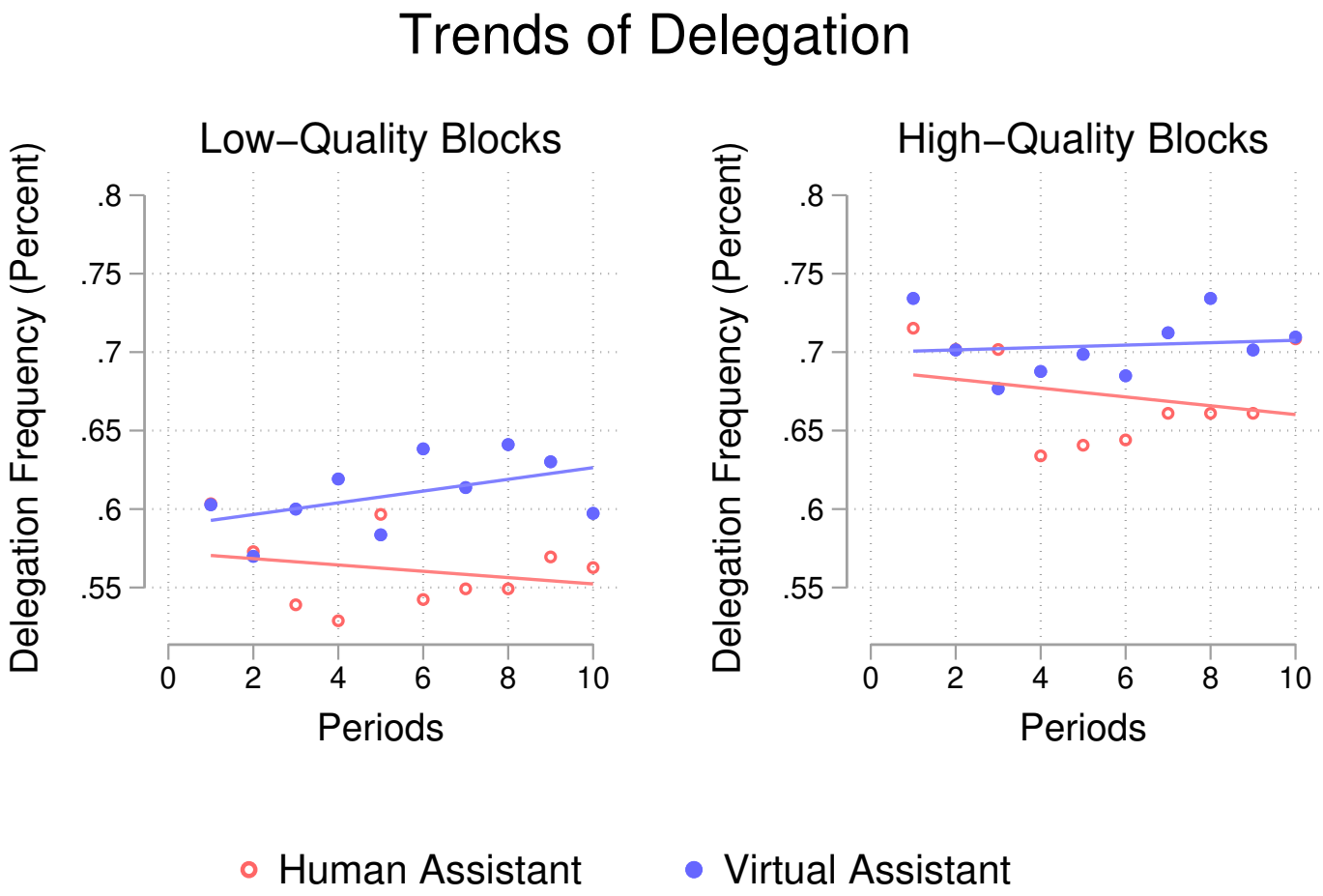
Manipulation Check: Human Assistant



Return

Trends of Delegation

1. No significant positive/negative trends. And no trend differences between Human and AI.



Does Virtual Asst Usage Predict Delegation?

- Post-experimental survey collected subjects' frequency of using virtual assistants.
- Variable "UsageFreq" is recoded such that higher value indicates higher frequency of use.

Screenshot of Unsuccessful Search

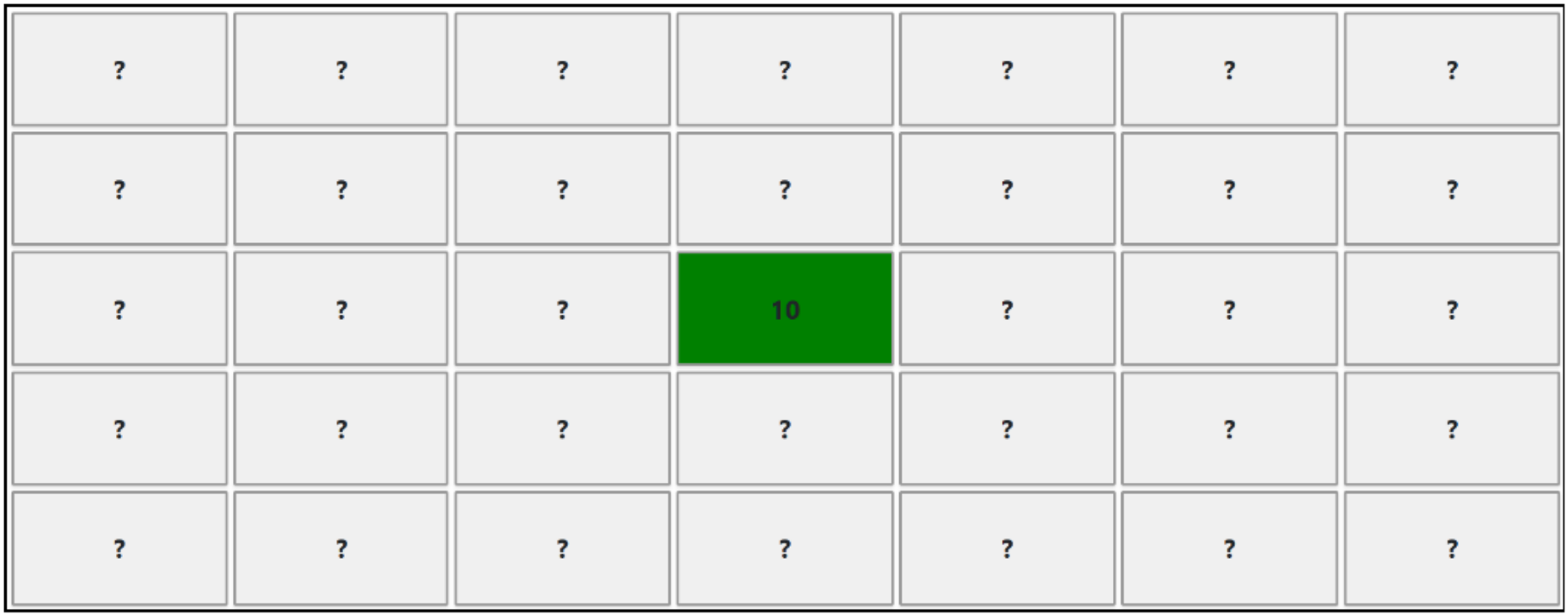


Figure: Unsuccesful Search: Only Green Card's position is revealed

Return

Screenshot of Successful Search

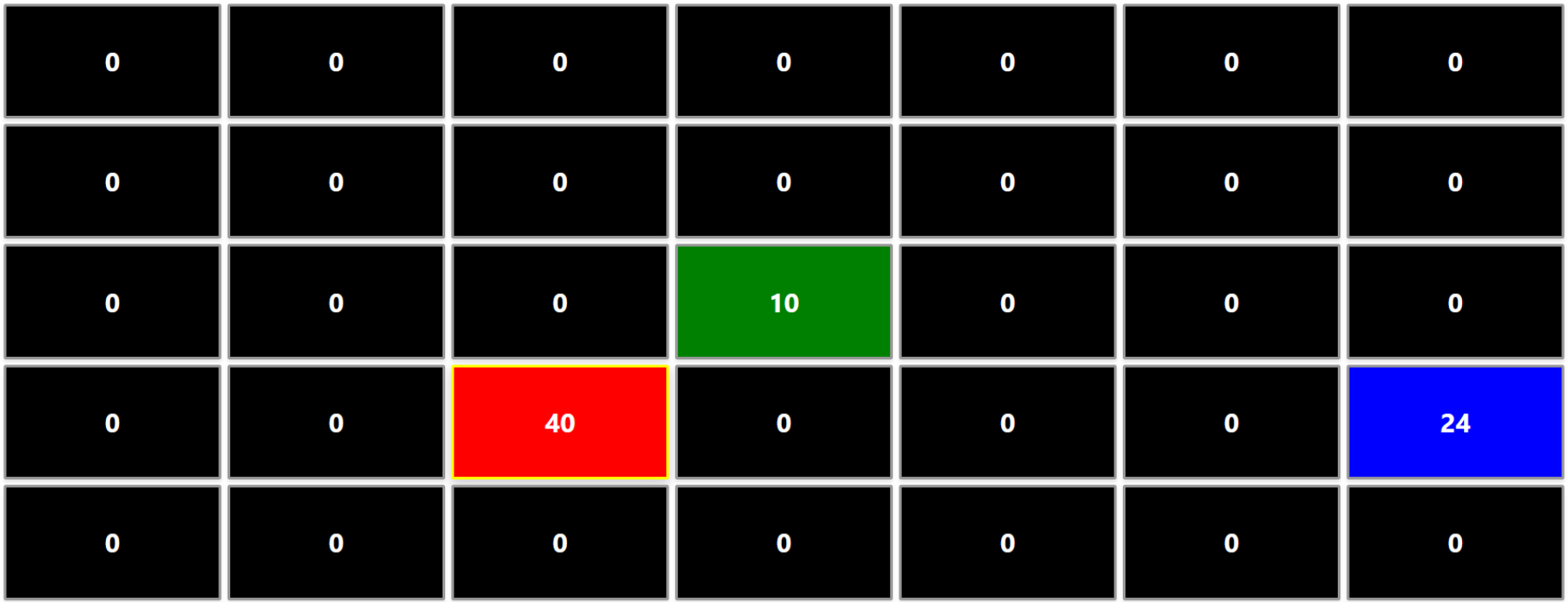


Figure: Succesful Search: All cards positions are revealed

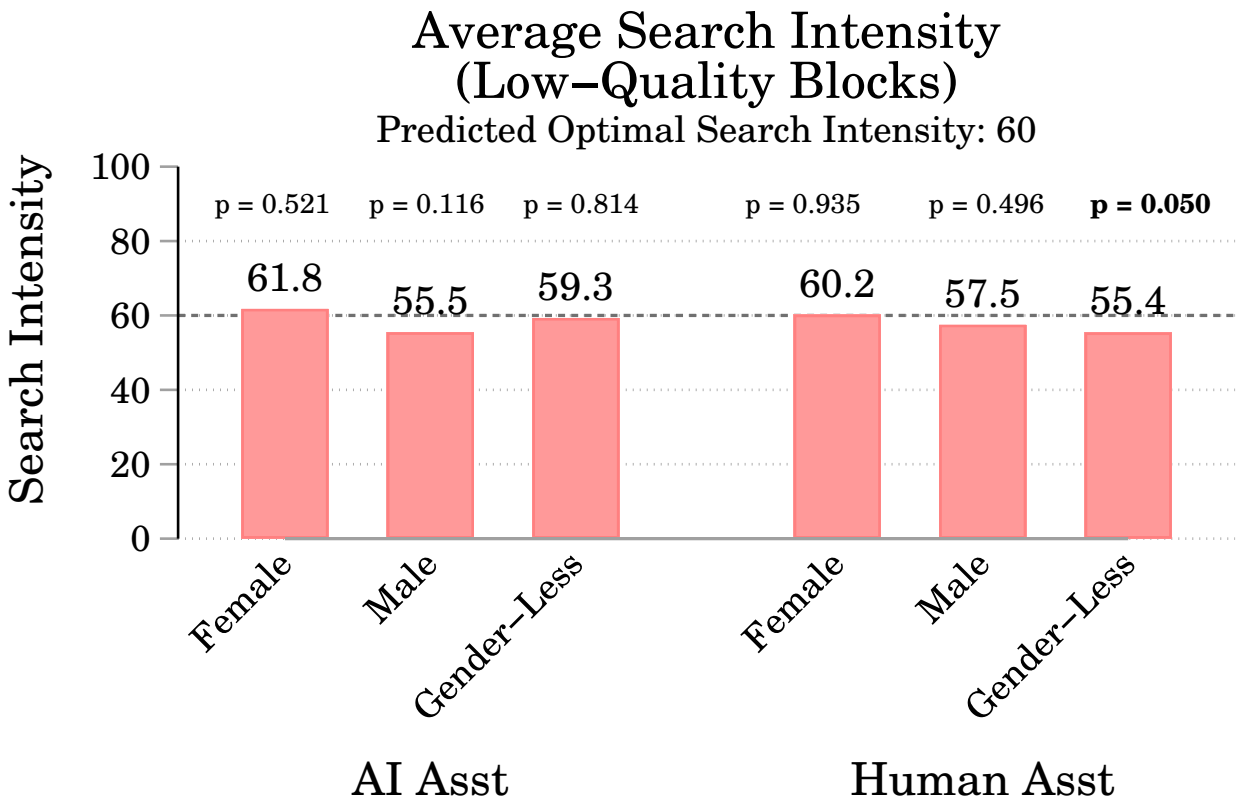
Return

Lab Sessions for Assistant Data

- Undergraduate subjects were invited to the Human Behavior Lab;
- Two blocks of information search varying in Quality;
- Each block: subjects chose intensity between 60% and 80% and repeated the search for 10 periods;
  - The sequence of *successful* and *unsuccessful* searches is fixed by seed for each intensity.
- Subjects chose a "pseudonym" name to represent themselves, indicating their gender;
- We selected subjects with the names "Jennifer" and "Charles" to construct the Human and AI assistants.
  - No deception, while making all assistants identical.

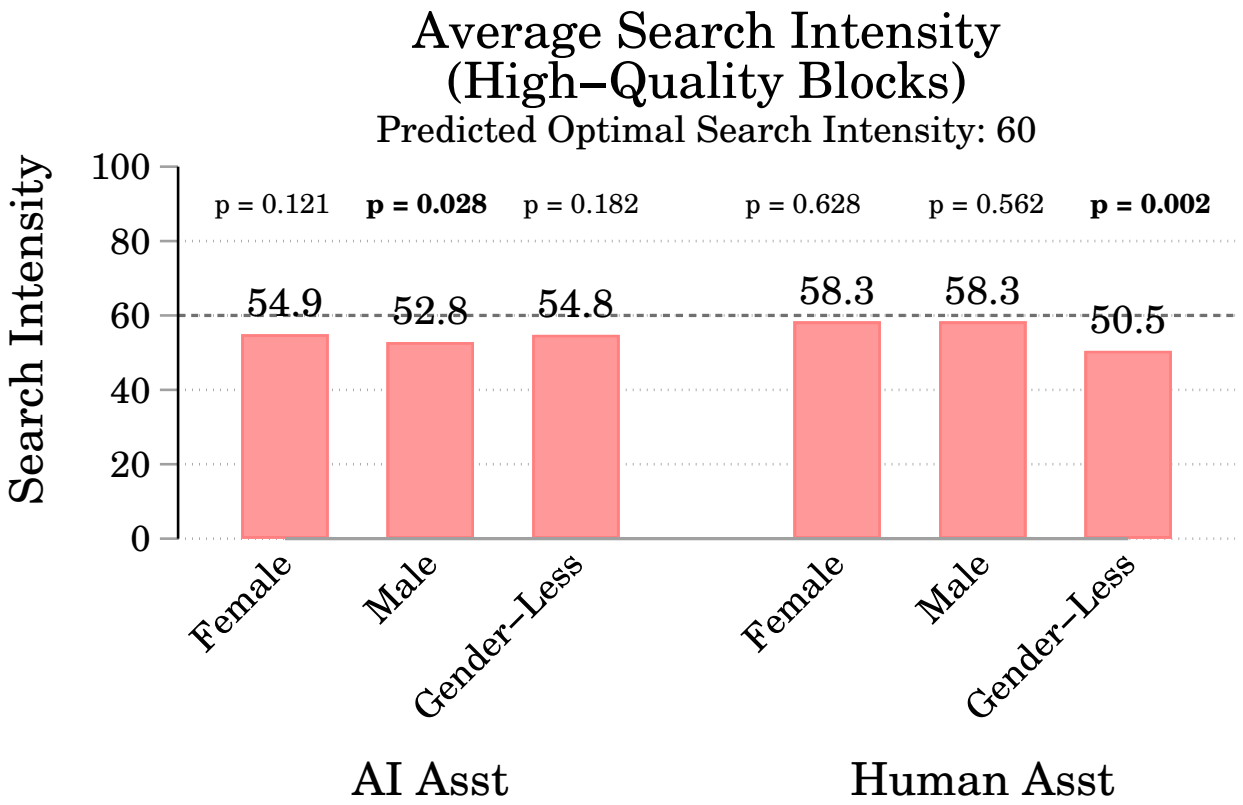
◂ Return

Search Intensity when Self-Search: Low-Quality Blocks



Return

Search Intensity when Self-Search: High-Quality Blocks



Return

Brief Implicit Association Test

1. Brief IAT: 2 blocks of tests, taking < 1 min to finish (Sriram et al., 2009);
2. Concept categories: male and female names;
3. Attribute categories: pleasant and unpleasant words;
4. D-score measures subjects' implicit association by measuring response times when classifying concepts and attributes;
5. In our case:
   - D-score > 0: association between **female** and pleasant words;
   - D-score ≤ 0: association between **male** and pleasant words;

Screenshot

Return

Screenshots of BIAT - 1

Instructions : Block 1

In this task, you will categorize items into groups as fast as you can. There are 2 blocks for which the instructions change. Please stay alert!
You are now in Block 1.

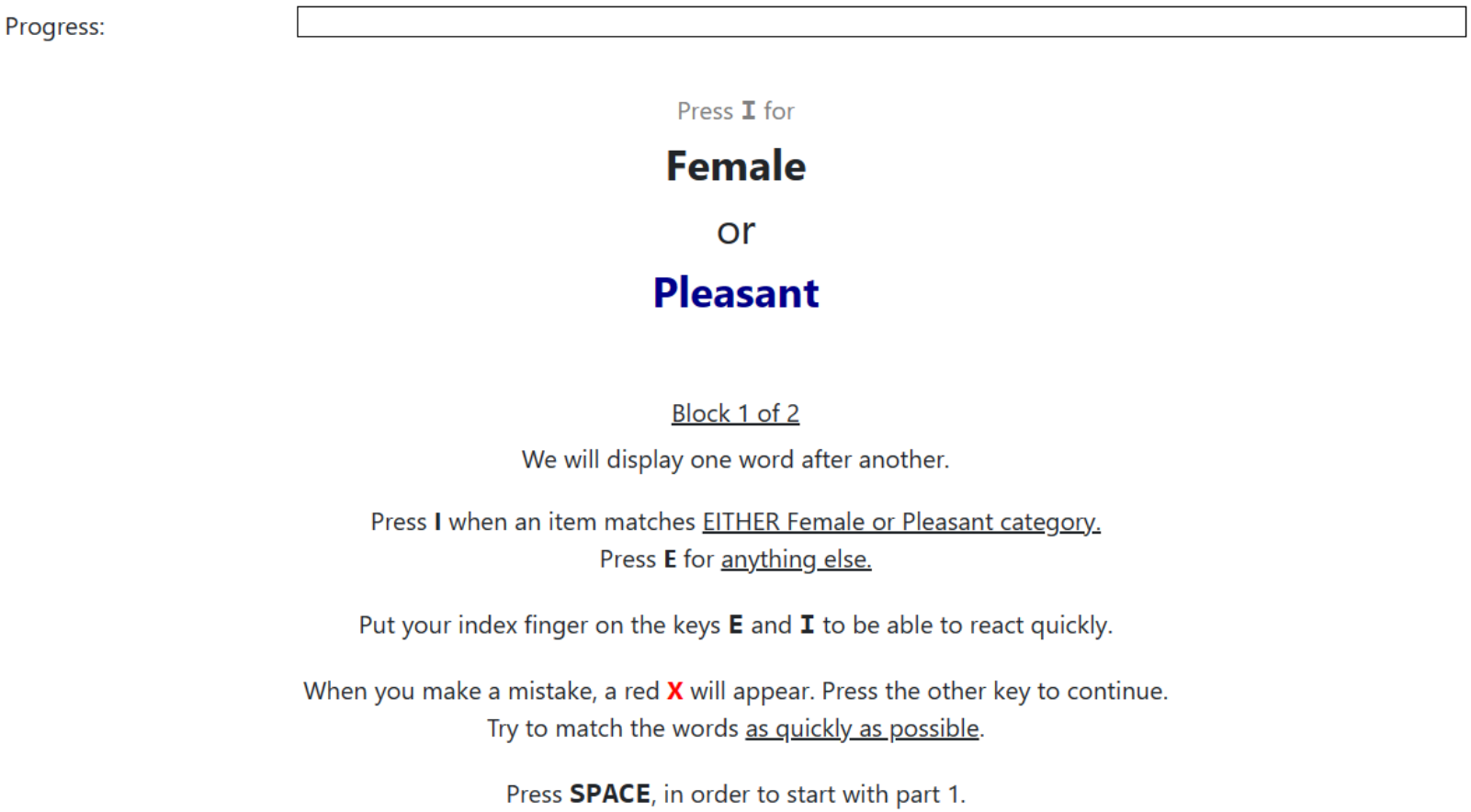
There are two categories. Their items are displayed below. Keep the two categories in your mind as you do the task.

Category	Word
Female	Amanda, Courtney, Heather, Melanie, Amber
Pleasant	good, loving, competent, smart, kind

Next

◂ Return

Screenshots of BIAT - 2



Screenshots of BIAT - 3

- In this example screenshot, subject needs to press "I" because the word in the middle is a female name.



Return

Warning Page

Warning

This study contains quality checks. You might be excluded at any point of the study if you do not pass quality checks or if you are not paying attention. So, please read the instructions carefully.

Next

◂ Return

ME of VA Gender on Choice of Female HA

	DV: Choice of Female Human Assistant					
	Male & Female Subjects		Female Subjects		Male Subjects	
	(1) Low Quality	(2) High Quality	(3) Low Quality	(4) High Quality	(5) Low Quality	(6) High Quality
Female vs. Gender-less	0.005	0.087	0.096	0.097	-0.058	0.070
	(0.065)	(0.066)	(0.109)	(0.101)	(0.079)	(0.089)
Male vs. Gender-less	0.074	0.045	0.300***	-0.004	-0.119	0.074
	(0.064)	(0.064)	(0.099)	(0.104)	(0.084)	(0.084)
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	633	633	264	264	356	356

◂ Return