**Supplemental Materials for**

**The Human Black-Box:**

**The Illusion of Understanding Human Better than Algorithmic Decision-making**

Table of Content

A: Experiments A-C ……...…………...………………..…………..………………………….... 2

B: Screening Procedure …...………………...………………..……………………….…………. 4

C: Method, Stimuli, and Additional Analyses for Experiment 2 .………………….………..…... 5

D: Method, Stimuli, and Additional Analyses for Experiment 3 ….…......................................... 7

E: Experiment 4 ……………………...……………………………………………………..….. 10

F: References ……….…………………………………...……………………………………... 13

A: Experiments A-C

Experiments A-C test transparency demands for human versus algorithmic decision-makers in the domains of criminal justice (A), recruiting (B), and healthcare (C).

**Method**

Respondents recruited from MTurk were randomly assigned to a 2(decision-maker: human, algorithm) between-subject design. We aimed to collect 100 responses per condition. The sample size was predetermined, and a sensitivity power analysis (Faul et al., 2009) indicated that the study had the power to detect a small-to-medium effect (*d*=.40) with a significance level α of .05 and a power (1-β) of .80. The final samples consisted of 201 respondents (93 females, 106 males, 2 did not indicate gender; age: *M*=37.35, *SD*=13.26) in Experiment A, 201 respondents (88 females, 110 males, 3 did not indicate gender; age: *M*=38.46, *SD*=13.69) in Experiment B, and 200 respondents (100 females, 98 males, 2 did not indicate gender; age: *M*=39.57, *SD*=12.92) in Experiment C. Respondents read the following.

*Experiment A.* Recently, a man was found driving a car that had been stolen. He was arrested, found guilty of stealing the car, and sentenced to prison. To determine the length of the sentence, a judge [computerized algorithm] evaluated the risk that the defendant would re-offend. The judge [algorithm] evaluated the defendant as high risk. Based on the judge’s [algorithm’s] assessment, the defendant was sentenced to five years in prison.

*Experiment B.* An increasing number of companies are turning to recorded video interviews to screen job applicants. These are interviews where the interviewer is not present when a candidate answers questions. A recorded video interview works as follows. Employers define a set of questions. Candidates answer the questions in front of a camera, record a video, and send it to the employer. Videos are then analyzed by a recruiter, who evaluates [an artificial intelligence system that uses an algorithm to evaluate] candidates. Based on the recruiter’s [algorithm’s] assessment, some candidates are rejected, and others move on in the recruiting process.

*Experiment C.* Modern medicine is increasingly recurring to diagnostic imaging techniques -- such as X-ray, MRI and PET/CT scans -- to diagnose a wide range of medical conditions (e.g., lesions, cancer, pneumonia, etc.). Diagnostic imaging often involves a radiologist who [an artificial intelligence algorithm that] examines the images for certain features in patients’ anatomical structure and makes a medical diagnosis.

Respondents indicated their demand for transparency on three items (Experiment A: *α*=.64; Experiment B: *α*=.77; Experiment C: *α*=.71), as illustrated for Experiment A:

The judge [algorithm] should have to explain exactly what information was considered for the risk assessment. (1=Strongly Disagree; 7=Strongly Agree).

The judge [algorithm] should have to explain exactly how the information considered was weighted for the risk-assessment. (1=Strongly Disagree; 7=Strongly Agree).

It is ok for the judge [algorithm] not to have to explain exactly how the defendant’s risk of reoffending was determined. (1=Strongly Disagree; 7=Strongly Agree; reversed coded).

**Results**

Respondents presented with an algorithmic decision-maker demanded significantly higher levels of transparency (Experiment A: *M*=5.79, *SD*=1.04; *M*=5.66, *SD*=1.18; Experiment C: *M*=5.77, *SD*=.88) than respondents presented with a human decision-maker (Experiment A: M=5.17, *SD*=1.09; *t*(199)=4.15, *p*<.001, *d*=.59; Experiment B: *M*=4.73, *SD*=1.36; *t*(199)=5.16, *p*<.001, *d*=.73; Experiment C: *M*=5.43, *SD*=.98; *t*(198)=2.52, *p*=.012, *d*=.36).

B: Screening Procedure

The following tasks were administered at the beginning of each study to screen for bots and avoid differential dropout. Participants who did not pass the screenings were redirected to the end of the survey without being assigned to conditions, thus did not affect the sample size count.

**Screening 1**

The purpose of this task was to screen for bots. Participants were provided with the instructions displayed below. The instructions were presented through a picture so they could not be processed by bots. This task also allowed to screen for inattentive respondents.

“Please click anywhere on this page at least five times. The number of times you

click will be recorded automatically, even if you do not see any change on this page.

This is an attention check. Failing this attention check will result in rejecting your hit. Thank you!”

Participants who failed this screening were redirected to the end of the survey. Participants who passed it were presented with an additional screening task described below.

**Screening 2**

The task below was included to screen out respondents unwilling to elaborate on the open-ended explanation task presented in the study, thus reducing the risk of differential dropout (Zhou & Fischbach, 2016).

**READ CAREFULLY!**

In this survey you will be asked some **open-ended questions** that require you to

provide answers in **writing**. Please **do not continue** if you do not want to answer questions that require writing. If you decide to continue, overall the survey will take about **3 to 5 mins**.”

Get out of the survey Continue the survey

Participants who chose to “get out of the survey” were automatically redirected to the end of the survey. Participants who chose to “continue the survey” were asked to engage in the writing task below to reinforce the expectation that the survey would involve open-ended questions that required writing.

Briefly describe the room you are currently in.

Only participants who successfully passed these screenings were randomly assigned to conditions.

C: Method, Stimuli, and Additional Analyses for Experiment 2

**Method**

Respondents were recruited from MTurk and randomly assigned to a 2 (self-understanding: high, low) by 2(decision-maker: human, algorithm) by 2(understanding: pre-explanation, post-explanation) mixed-design, with repeated measures on understanding. We aimed to collect 100 responses per condition. The sample size was predetermined, and a sensitivity power analysis (Faul et al. 2009) indicated that the study had the power to detect a small effect (*ηp*2=.01) with a significance level α of .05 and a power (1-β) of .80. The final sample consisted of 400 respondents (222 females, 175 males, 3 did not indicate gender; age: *M*=39.17, *SD*=12.41) who passed the initial screening (see Supplemental Material B) and completed the survey. There was no differential dropout across conditions (*χ2*(3)=4.89, *p*=.180).

**Procedure**

We adapted the stimuli used in Experiment 1A. Specifically, participants read the following. In the United States, a criminal offender who has been sentenced to prison can become eligible for parole after serving part of the given sentence. When criminal offenders are paroled, they are released from prison and serve the remaining of the sentence in the community under supervision conditions: Decisions about parole entail an evaluation of the risk that a defendant will re-offend if released.

*Self-understanding manipulation:* How easy [difficult] is it to evaluate the risk that a defendant will re-offend? It turns out that these evaluations are rather easy [very difficult] to make. Recent research shows that even ordinary people are pretty good [that ordinary people are very bad] at evaluating whether a defendant has a high or low risk to re-offend.

*Self-understanding:* Do you understand how you would evaluate the risk that a defendant will re-offend? (1=Do not understand at all; 7=Completely understand)

*Decision-maker manipulation:* In many jurisdictions, a judge [an algorithm] evaluates the risk that a defendant will re-offend if released.

*Understanding before explanation:* Do you understand how a judge [an algorithm] evaluates the risk that a defendant will re-offend if released? (1=Do not understand at all; 7=Completely understand).

*Explanation manipulation:* If you know it, please explain in detail, step by step, the process used by a judge [an algorithm] to evaluate the risk that a defendant will re-offend if released. If there are aspects that you don’t know or cannot explain, write “GAP” in your description at that point.

*Understanding after explanation:* In light of the task you just completed, you might have changed your mind about how deeply and thoroughly you understand how a judge [an algorithm] evaluates the risk that a defendant will re-offend. Next, please re-assess your understanding.

Do you understand how a judge [an algorithm] evaluates the risk that a defendant will re-offend if released? (1=Do not understand at all; 7=Completely understand).

**Additional Analyses**

A 2(decision-maker: human, algorithm) by 2(self-understanding: high, low) by 2(understanding: before explanation, after explanation) mixed ANOVA with mixed-design ANOVA with repeated measures on understanding revealed no significant main effects of decision-maker (*F*(1,396)=1.25, *p*=.265, *ηp2*=.003) or self-understanding (*F*(1,396)=2.48, *p*=.116, *ηp2*=.01), but a significant main effect of understanding (*F*(1,396)=102.22, *p*<.001, *ηp2*=.21). The two-way interactions decision-maker × self-understanding (*F*(1,396)=1.37, *p*=.242, *ηp2*=.003) and decision-maker × understanding (*F*(1,396)=.397, *p*=.529, *ηp2*=.001) were not significant, but the interaction self-understanding × understanding (*F*(1,396)=4.27, *p*=.039, *ηp2*=.01) was. The predicted three-way interaction was significant (*F*(1,396)=10.06, *p*=.002, *ηp2*=.03).

*Planned comparisons before explanation.* In the high self-understanding conditions, respondents’ indicated higher understanding of the judge’s decision-making process (*M*=4.64, *SD*=1.90) than of the algorithm’s (*M*=3.96, *SD*=1.97; *t*(396)=2.59, *p*=.010, *d*=.37). In the low self-understanding conditions, understanding of the decision-making process did not differ between decision-makers (*Mjudge*=3.76, *SDjudge*=1.69; *Malgorithm*=3.99, *SDalgorithm*=1.78; *t*(396)=.89, *p*=.373, *d*=.12). Viewed differently, manipulating respondents’ own sense of understanding significantly influenced their understanding of the judge’s decision-making process (*t*(396)=3.39, *p*<.001, *d*=.48), but not of the algorithm’s (*t*(396)=.13, *p*=.899, *d*=.02).

*Planned comparisons after explanation.* Respondents’ understanding of the decision-making process after the explanation task did not differ across conditions (judge high self-understanding: *M*=3.26, *SD*=1.74; algorithm high self-understanding: *M*=3.23, *SD*=1.78; judge low self-understanding: *M*=3.29, *SD*=1.76; algorithm low self-understanding: *M*=3.08, *SD*=1.54; *F*(3, 396)=.32, *p*=.811, *ηp2*=.002).

*IOED.* We computed an IOED score by subtracting the understanding ratings for the judge/algorithm after the explanation task from the understanding ratings before the explanation task. A 2(decision-maker: human, algorithm) by 2(self-understanding: high, low) between-subjects ANOVA on such IOED score revealed no significant main effect of decision-maker (*F*(1,396)=.397, *p*=.529, *ηp2*=.001), a significant main effect of the self-understanding manipulation (*F*(1,396)=4.27, *p*=.039, *ηp2*=.01), and a significant interaction (*F*(1,396)=10.06, *p*=.002, *ηp2*=.03). In the high self-understanding conditions, respondents displayed a significantly larger illusion of understanding for the judge (*M*=1.38, *SD*=1.89) than for the algorithm (*M*=.72, *SD*=1.75; *t*(396)=2.65, *p*=.008, *d*=.38), thus replicating Experiment 1A. In the low self-understanding conditions, however, no significant difference emerged (*Mjudge*=.48, *SDjudge*=1.83; *Malgorithm*=.91, *SDalgorithm*=1.41; *t*(396)=1.83, *p*=.069, *d*=.25). From a different angle, manipulating respondents’ own sense of understanding significantly influenced the illusion of understanding the judge (*t*(396)=3.72, *p*<.001, *d*=.52), but not the algorithm (*t*(396)=.78, *p*=.437, *d*=.11).

**D: Method, Stimuli, and Additional Analyses for Experiment 3**

**Method**

Respondents recruited from MTurk were randomly assigned to a 2(decision-maker: human, algorithm) by 2(dissimilarity: control, dissimilar) by 2(understanding: pre-explanation, post-explanation) mixed-design, with repeated measures on understanding. We aimed to collect 100 responses per condition. The sample size was predetermined, and a sensitivity power analysis (Faul et al., 2009) indicated that the study had the power to detect a small effect (*ηp*2=.01) with a significance level α of .05 and a power (1-β) of .80. The final sample consisted of 400 respondents (200 females, 197 males, 3 did not indicate gender; age: *M*=39.92, *SD*=12.78) who passed the initial screening (see Supplemental Material B) and completed the survey. There was no differential dropout across conditions (*χ2*(3)=3.05, *p*=.385).

**Procedure**

We adapted the stimuli used in Experiment 1C.

*Decision-Maker Manipulation:* Osteoarthritis is a very common condition that affects millions of people worldwide. It occurs when the protective cartilage that cushions the ends of our bones wears down over time. It can affect any joint in the body, and it’s most likely to affect the joints that we use most in everyday life, such as the joints of the hands, knees, feet, elbows and neck. To diagnose osteoarthritis, a technician takes MRI images of the joints. The images are then examined by a radiologist [an artificial intelligence algorithm].

*Dissimilarity Manipulation:* While you may have a few trivial things in common with a radiologist [an artificial intelligence algorithm], you are likely very different from a radiologist [an artificial intelligence algorithm] in many fundamental ways. Think about the fundamental ways in which you are DIFFERENT from a radiologist [an artificial intelligence algorithm]. In the space below, please elaborate on fundamental things that make you DIFFERENT from a radiologist [an artificial intelligence algorithm].

*Explanation manipulation:* If you know it, please explain in detail the process used by a radiologist [an artificial intelligence algorithm] to examine MRI images to diagnose osteoarthritis. If there are aspects that you don’t know or cannot explain, write “GAP” in your description at that point.

*Understanding before explanation:* Do you understand how a radiologist [an artificial intelligence algorithm] examines MRI images to diagnose osteoarthritis? (1=Do not understand at all; 7=Completely understand).

*Understanding after explanation:* In light of the task you just completed, you might have changed your mind about how well you understand how a radiologist [an artificial intelligence algorithm] examines MRI images to diagnose osteoarthritis. Next, please re-assess your understanding. Do you understand how a radiologist [an artificial intelligence algorithm] examines MRI images to diagnose osteoarthritis? (1=Do not understand at all; 7=Completely understand).

**Manipulation Check**

To verify the effectiveness of the dissimilarity manipulation, we recruited a sample of respondents from the same population. Respondents were assigned to the same 2(decision-maker: human, algorithm) by 2(dissimilarity: control, dissimilar) between-subjects design as in the main experiment. We aimed to collect 50 responses per condition. The sample size was predetermined, and a sensitivity power analysis (Faul et al., 2009) indicated that the study had the power to detect a small effect (*ηp*2=.04) with a significance level α of .05 and a power (1-β) of .80. The final sample consisted of 200 respondents (99 females, 99 males, 2 did not indicate gender; age: *M*=42.61, *SD*=13.48) who passed the initial screening (see Supplemental Material B) and completed the survey. There was no differential dropout across conditions (*χ2*(3)=2.68, p=.444).

Respondents rated their perceived similarity to the decision-maker (human or algorithm) by indicating their agreement with two statements (overall, there are things that make a radiologist [algorithm] similar to me; in general, I have some characteristics in common with a radiologist [algorithm]; 1=Disagree; 7=Agree; *r*=.835, *p*<.001).

A 2(decision-maker: human, algorithm) by 2(dissimilarity: control, dissimilar) between-subjects ANOVA on perceived similarity revealed a main effect of decision-maker (*F*(1,196)=11.52, *p*<.001, *ηp2*=.06), a marginally significant main effect of dissimilarity (*F*(1,196)=3.01, *p*=.084, *ηp2*=.02), and a significant interaction (*F*(1,196)=5.62, *p*=.019, *ηp2*=.03). In the control conditions, respondents indicated higher similarity to a radiologist (*M*=4.59, *SD*=1.51) than to an algorithm (*M*=3.28, *SD*=1.67; *t*(196)=4.15, *p*<.001, *d*=.82). No difference emerged in the dissimilar conditions (*Mradiologist*=3.66, *SDradiologist*=1.44; *Malgorithm*=3.42, *SDalgorithm*=1.74, *t*(196)=.71, *p*=.477, *d*=.15). Specifically, the dissimilarity manipulation reduced perceived similarity significantly for the radiologist (*t*(196)=2.85, *p*=.005, *d*=.58), but did not change perceived similarity for the algorithm (*t*(196)=.46, *p*=.647, *d*=.09), indicating that the manipulation worked as intended.

**Additional Analyses**

A 2(decision-maker: human, algorithm) by 2(dissimilarity: control, dissimilar) by 2(understanding: before explanation, after explanation) mixed-design ANOVA with repeated measures on understanding revealed a marginally significant main effect of decision-maker (*F*(1,396)=2.79, *p*=.096, *ηp2*=.01), no significant main effect of dissimilarity (*F*(1,396)=1.23, *p*=.268, *ηp2*=.003), and a significant main effect of understanding (*F*(1,396)=94.50, *p*<.001, *ηp2*=.19). The two-way interactions decision-maker × dissimilarity (*F*(1,396)=5.41, *p*=.021, *ηp2*=.01) as well as dissimilarity × understanding (*F*(1,396)=4.54, *p*=.034, *ηp2*=.01) were significant, while the interaction decision-maker × understanding was not *F*(1,396)=.93, *p*=.335, *ηp2*=.002). The predicted three-way interaction was significant (*F*(1,396)=10.13, *p*=.002, *ηp2*=.03).

*Planned comparisons before explanation.* In the control conditions, participants indicated higher understanding for the radiologist (*M*=4.17, *SD*=1.76) than for the algorithm (*M*=3.23, *SD*=1.68; *t*(396)=3.66, *p*<.001, *d*=.52). In the dissimilarity conditions, sense of understanding did not differ between decision-makers (*Mradiologist*=3.25, *SDradiologist*=1.91; *Malgorithm*=3.51, *SDalgorithm*=1.84, *t*(396)=1.00, *p*=.316, *d*=.14). Viewed differently, the dissimilarity manipulation reduced the sense of understanding significantly for the radiologist (*t*(396)=3.65, *p*<.001, *d*=.51), but not for the algorithm (*t*(396)=1.05, *p*=.295, *d*=.15).

*Planned comparisons after explanation.* Respondents’ understanding of the decision-making process after the explanation task did not differ across conditions (radiologist control: *M*=3.11, *SD*=1.81; algorithm control: *M*=2.72, *SD*=1.72; radiologist dissimilar: *M*=2.89, *SD*=1.78; algorithm dissimilar: *M*=2.86, *SD*=1.70; *F*(3, 396)=.83, *p*=.480, *ηp2*=.01).

*IOED.* We computed an IOED score by subtracting the understanding ratings for the radiologist/algorithm after the explanation task from the understanding ratings before the explanation task. A 2(decision-maker: human, algorithm) by 2(dissimilarity: control, dissimilar) between-subjects ANOVA on such IOED score revealed no significant main effect of decision-maker (*F*(1,396)=.930, *p*=.335, *ηp2*=.002), a significant main effect of dissimilarity (*F*(1,396)=4.54, *p*=.034, *ηp2*=.01), and a significant interaction (*F*(1,396)=10.13, *p*=.002, *ηp2*=.03). In the control conditions, respondents displayed a significantly larger illusion of understanding for the radiologist (*M*=1.06, *SD*=1.54) than for the algorithm (*M*=.51, *SD*=1.18; *t*(396)=2.93, *p*=.004, *d*=.42), thus replicating Experiment 1C. In the dissimilarity conditions, however, no difference emerged (*Mradiologist*=.36, *SDradiologist*=1.39; *Malgorithm*=.65, *SDalgorithm*=1.15, *t*(396)=1.57, *p*=.117, *d*=.22). Specifically, the dissimilarity manipulation reduced the illusion of understanding the radiologist (*t*(396)=3.81, *p*<.001, *d*=.53), but not the algorithm (*t*(396)=.74, *p*=.463, *d*=.11).

**E: Experiment 4**

Experiment 4 aims to provide further converging evidence for our proposed projection mechanism via moderation by similarity in the recruiting domain, and test whether illusory understanding fosters greater trust in a decision made by a human as opposed to an algorithm. In doing so, it also aims to replicate the results of Experiment 1B.

**Method**

Respondents recruited from MTurk were randomly assigned to three between-subject conditions (decision-maker: recruiter, recruiter dissimilar, algorithm). We aimed to collect 100 responses per condition. The sample size was predetermined, and a sensitivity power analysis (Faul et al., 2009) indicated that the study had the power to detect a small effect (*η*2=.03) with a significance level α of .05 and a power (1-β) of .80. The final sample consisted of 300 respondents (164 females, 134 males, 2 did not indicate gender; age: *M*=39.41, *SD*=13.42), who passed the initial screening (see Supplemental Material B) and completed the survey. There was no differential dropout across conditions (*χ2*(2)=2.74, *p*=.255).

**Procedure**

We adapted the stimuli used in Experiment 1B.

*Decision-Maker Manipulation:* Companies are turning to recorded video interviews to screen job applicants. Candidates answer questions in front of a camera, record a video, and send it to the employer. A recruiter [an algorithm] then reviews the video and evaluates the candidate.

*Dissimilarity Manipulation:* While you may have a few things in common with a recruiter [an algorithm], you are likely very different from a recruiter [an algorithm] in many fundamental ways. Think about the fundamental ways in which you are DIFFERENT from a recruiter [an algorithm]. In the space below, please elaborate on fundamental things that make you DIFFERENT from a recruiter [an algorithm].

Note that only participants in the recruiter dissimilar condition completed the dissimilarity manipulation before rating understanding and trust. In the recruiter and algorithm conditions respondents completed the dissimilarity manipulation only after rating understanding and trust. Thus, this task was irrelevant in these conditions (because it was presented after the dependent variables), yet it ensured that all conditions were similar in terms of the effort required to complete the study, thus reducing the risk of differential dropout (Zhou & Fischbach, 2016).

*Measure of Understanding:* Do you understand how a recruiter [algorithm] reviews a video to evaluate a candidate? (1=Do not understand at all; 7=Completely understand).

*Measure of Trust:* Now imagine that you applied for a job at a company where a recruiter [an algorithm] screens candidates based on recorded video interviews, and you did not pass the screening. How much would you trust the recruiter’s [algorithm’s] decision? (1=Not at all; 7=Very much).

**Manipulation Check**

To verify the effectiveness of the dissimilarity manipulation, we recruited a sample of respondents from the same population. Respondents were assigned to the same three between-subject conditions (decision-maker: recruiter, recruiter dissimilar, algorithm) as in the main experiment. We aimed to collect 50 responses per condition. The sample size was predetermined, and a sensitivity power analysis (Faul et al., 2009) indicated that the study had the power to detect a moderate effect (*η*2=.06) with a significance level α of .05 and a power (1-β) of .80. The final sample consisted of 150 respondents (77 females, 72 males, 1 did not indicate gender; age: *M*=42.15, *SD*=13.39) who passed the initial screening (see Supplemental Material B) and completed the survey. There was no differential dropout across conditions (*χ2*(2)=2.68, *p*=.263).

Respondents rated their perceived similarity to the decision-maker (human or algorithm) by indicating their agreement with two statements (overall, there are things that make a recruiter [algorithm] similar to me; in general, I have some characteristics in common with a recruiter [algorithm]; 1=Disagree; 7=Agree; *r*=.939, *p*<.001).

A one-way ANOVA on perceived similarity revealed a significant effect (*F*(2,147)=18.51, *p*<.001, *η2*=.20). Participants indicated greater perceived similarity with a recruiter (*M*=4.86, *SD*=1.11) compared to both a dissimilar recruiter (*M*=3.68, *SD*=1.78; *t*(147)=3.68, *p*<.001, *d*=.74) and an algorithm (*M*=2.99, *SD*=1.80; *t*(147)=6.02, *p*<.001, *d*=1.18). Perceived similarity with a dissimilar recruiter was higher than with an algorithm (*t*(147)=2.16, *p*=.033, *d*=.44).

**Results**

*Understanding.* A one-way ANOVA on understanding revealed a significant effect (*F*(2,297)=21.18, *p*<.001, *η2*=.13). Participants indicated a greater understanding of how a recruiter analyzes a video interview (*M*=5.61, *SD*=1.49) compared to both a dissimilar recruiter (*M*=4.58, *SD*=1.59; *t*(297)=4.25, *p*<.001, *d*=.61) and an algorithm (*M*=4.09, *SD*=1.93; *t*(297)=6.41, *p*<.001, *d*=.90). Understanding for the dissimilar recruiter was higher than that for the algorithm (*t*(297)=2.09, *p*=.037, *d*=.29).

*Trust.* A one-way ANOVA on trust revealed a significant main effect of decision-maker (*F*(2,297)=17.41, *p*<.001, *η2*=.11). Participants indicated higher trust in a recruiter’s decision (*M*=4.05, *SD*=1.44) compared to both a dissimilar recruiter’s decision (*M*=3.63, *SD*=1.36; *t*(297)=2.03, *p*=.043, *d*=.29) and an algorithm’s decision (*M*=2.88, *SD*=1.51; *t*(297)=5.80, *p*<.001, *d*=.82). Trust in the dissimilar recruiter’ decision was higher than in the algorithm’s decision (*t*(297)=3.74, *p*<.001, *d*=.53).

**Table L–1**

*Descriptive statistics in Experiment 4*

|  |  |  |  |
| --- | --- | --- | --- |
| Condition | N | UnderstandingM (SD) | TrustM (SD) |
| Recruiter | 97 | 5.61 (1.49) | 4.05 (1.44) |
| Recruiter Dissimilar | 98 | 4.58 (1.59) | 3.63 (1.36) |
| Algorithm | 105 | 4.09 (1.93) | 2.88 (1.51) |

*Mediation.*To test whether sense of understanding drove trust in the decision, we conducted a mediation analysis (decision-maker🡪understanding🡪trust) using the Process Macro Model 4 (Hayes, 2018) with 10,000 bootstraps resamples. We specified a 3-level, multi categorical independent variable (i.e., decision-maker: recruiter, recruiter dissimilar, algorithm), with the recruiter condition as the reference point in the tests of group differences. Regression coefficients are presented in Figure F-1. Each coefficient represents the contrast test against the recruiter condition. Comparing the recruiter to the algorithm yields a significant indirect effect on trust via understanding (*b*=-.44, *SE*=.10, CI95% [-.65;-.26]). However, after adjusting for the indirect effect, the direct effect of recruiter versus algorithm on trust remains significant (*b*=-.74, *SE*=.20, CI95% [-1.14;-.33]), indicating partial mediation. Comparing the recruiter to the dissimilar recruiter yields a significant indirect effect on trust via understanding as well (*b*=-.30, *SE*=.08, CI95% [-.48;-.15]). After adjusting for the indirect effect, the direct effect of recruiter versus dissimilar recruiter on trust is no longer significant (*b*=-.12, *SE*=.20, CI95% [-.52;.27]), indicating full mediation. Overall, these results suggest that the perception of similarity drove differences in understanding that were then associated with different levels of trust in the decision.

**Figure F–1.**

*Mediation Model in Experiment 4*

Sense of Understanding

.29\*\*\*

X1: -1.52\*\*\*

X2:-1.03\*\*\*

Trust

Decision-maker

X1: *-*.74\*\*\* (-1.18\*\*\*)

X2:-.12 (-.42\*)

*Note. X1: recruiter vs. algorithm; X2: recruiter vs. recruiter dissimilar; \*p<.05; \*\*p<.01; \*\*\*p<.001; total effects presented in brackets.*

**F: References**

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. (2009). Statistical power analyses using G\* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149-1160.

Hayes, A. (2018). Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-based Approach. The Guilford Press, New York, NY.

Zhou, H., & Fishbach, A. (2016). The pitfall of experimenting on the web: How unattended selective attrition leads to surprising (yet false) research conclusions. *Journal of Personality and Social Psychology*, *111*(4), 493–504