

1 **Recalled AI Experiences Shift Evaluations of AI but Not of Humans or Experts**

2 **Supplementary Information**

3
4 Kasumi Dan ^{1,2*}, Takahiro Hoshino ^{2,3}

5
6 ¹ Graduate School of Economics, Keio University, Tokyo, Japan

7 ² RIKEN Center for Advanced Intelligence Project (AIP), RIKEN (National Institute of Physical and
8 Chemical Research), Tokyo, Japan

9 ³ Department of Economics, Keio University, Tokyo, Japan

10
11 *Corresponding author: Kasumi Dan

12 Email: kasumi22389@keio.jp

13
14 Reference numbers correspond to the reference list in the main manuscript.

15
16 Supplementary Table S1. Priming Stimuli (3 conditions)

Condition	Stimulus passage
Satisfying AI experience priming	Please describe three specific experiences or aspects of your use of generative AI that you felt positive about, or that you found satisfying, such as instances where it was useful, helpful, or impressive.
Dissatisfying AI experience priming	Please describe three specific experiences or aspects of your use of generative AI that you felt negative about, or that you found unsatisfying, such as instances where it was inconvenient, not useful, or disappointing.
Neutral AI experience priming	Please describe three specific ways in which you have used generative AI so far.

17
18 Priming instructions used in the three AI experience recall conditions (satisfying, dissatisfying, and
19 neutral). Participants were asked to describe three aspects of their prior use of generative AI
20 corresponding to the assigned condition.

21 Supplementary Table S2. Measurement instruments and Cronbach's alphas

Latent construct and items	α
<i>Trust [62, 63]</i>	AI:0.85
	human:0.81
<ul style="list-style-type: none"> ● I generally trust (AI/ other people/ experts). ● I tend to count upon (AI/ other people/ experts). ● I generally have faith in (AI/ humanity / experts). 	expert:0.88
<i>Perceived Competence [64]</i>	AI:0.84
	human:0.80
In general, (AI / human beings /experts) are...	expert:0.80
<ul style="list-style-type: none"> ● not very competent — very competent ● not very efficient — very efficient ● not very intelligent — very intelligent 	
<i>Willingness to Rely (AI-appropriate tasks)</i>	AI:0.82
	human:0.82
<ul style="list-style-type: none"> ● When I want knowledge, information, or an answer to a question, I would like to ask (an AI / a human/ an expert). ● When I need help creating text or documents (e.g., emails, reports, summaries), I would like to get help from (an AI / a human / an expert). ● When I want small forms of everyday support to improve my daily life (e.g., meal ideas, travel planning, light health advice), I would like to rely on (an AI / a human/ an expert). ● When I need support with studying or learning (e.g., clear explanations, practice questions, study plans), I would like to be helped by (an AI / a human/ an expert). 	expert:0.83

<i>Willingness to Rely (Human-appropriate tasks)</i>	AI:0.80 human:0.74 expert:0.70
<ul style="list-style-type: none"> ● When I want to talk about personal concerns or sort out my feelings, I would like to turn to (an AI / a human/ an expert). ● When I seek advice for major life decisions (e.g., career choices, job changes, where to live), I would like to consult (an AI / a human/ an expert). ● When I want to see creative works such as stories, artwork, or music, I would like (an AI / a human/ an expert) to create them. ● When I want to think through a complex topic with no clear answer or with divided opinions, I would like to discuss it with (an AI / a human/ an expert). 	
<i>AI Continuance Intention [65, 66]</i>	0.90
<ul style="list-style-type: none"> ● If I could, I would like to continue my use of generative AI tools. ● I intend to use generative AI tools in the next 6 months. ● I predict I would use generative AI tools in the next 5 years. 	
<i>Willingness to Pay for a Paid AI Plan [67]</i>	0.87
<ul style="list-style-type: none"> ● I intend to pay for premium generative AI tools. ● It is worthwhile to subscribe to premium generative AI tools. 	

22
23
24
25
26
27

Measurement items for each construct and corresponding internal consistency reliability (Cronbach's α). All items were rated on 7-point Likert scales. Target terms (AI, humans, or experts) were substituted as appropriate across conditions. Reliability coefficients are reported separately for each evaluation target where applicable.

28 Supplementary Table S3. Items and one-sample t-test results from the pilot study

Item	M	SD	t(59)	p	Classification
When you need help creating text or documents (e.g., emails, reports, summaries), do you think that assistance is typically provided by humans or by AI?	5.35	1.53	6.85	< 0.001	AI-appropriate
When you need support with studying or learning (e.g., clear explanations, practice questions, study plans), do you think such support is typically provided by humans or by AI?	4.82	1.52	4.15	< 0.001	AI-appropriate
When you want small forms of everyday support to improve your daily life (e.g., meal ideas, travel planning, light health advice), do you think such support is typically provided by humans or by AI?	4.68	1.41	3.76	< 0.001	AI-appropriate
When you want knowledge, information, or answers to questions, do you think they are typically taught or provided by humans or by AI?	4.58	1.62	2.79	0.007	AI-appropriate
When you want to discuss a complex topic with no clear answer or with divided opinions, do you think humans or AI typically serve as the discussion partner?	3.08	1.60	-4.45	< 0.001	Human-appropriate
When you want to talk about personal concerns or sort out your feelings, do you think humans or AI typically serve as the conversation partner?	2.95	1.63	-4.99	< 0.001	Human-appropriate
When it comes to advice for major life decisions (e.g., career choices, job changes, where to live), do you think it is typically provided by humans or by AI?	2.62	1.22	-8.77	< 0.001	Human-appropriate
When you want creative works such as stories, artwork, or music, do you think they are typically created by humans or by AI?	2.15	1.40	-10.24	< 0.001	Human-appropriate

29

30 Pilot study results validating task classification (N = 60). Participants rated whether each task is
 31 typically performed by humans (lower values) or AI (higher values) on a 7-point scale. One-sample t-
 32 tests compared each mean against the scale midpoint (4). Tasks with means significantly above the
 33 midpoint were classified as AI-appropriate; those significantly below were classified as human-
 34 appropriate.