Supplementary Information

Title: Rapid, precise, and reliable measurement of delay discounting using a Bayesian learning algorithm

Authors:

Woo-Young Ahn[1,2], Hairong Gu[2], Yitong Shen[3], Nathaniel Haines[2], Hunter A. Hahn[2], Julie E. Teater[4], Jay I. Myung[2], and Mark A. Pitt[2]

[1] Department of Psychology, Seoul National University, Seoul, Korea

[2] Department of Psychology, The Ohio State University, Columbus, OH

[3] Department of Psychiatry, Indiana University-Purdue University Indianapolis, Indianapolis, IN

[4] Department of Psychiatry and Behavioral Health, The Ohio State University, Columbus, OH

Correspondence:

Woo-Young Ahn, Ph.D.

Department of Psychology

Seoul National University

Seoul, Korea 08826

Tel: +82-2-880-2538, Fax: +82-2-877-6428. Email: wahn55@snu.ac.kr

Number of Figures/Tables in Supplementary Information: 11 / 0

Figure S1. Within-visit test-retest reliability of temporal discounting rates across two ADO (A-B) sessions and SC (C-D) sessions (Experiment 1, college students, with all 42 trials per session). Error bars on each point are Bayesian credible intervals defined in terms of ± 1 standard deviation of each participant's posterior distribution of the discounting rate parameter at the end of the session. CCC: Concordance Correlation Coefficient.
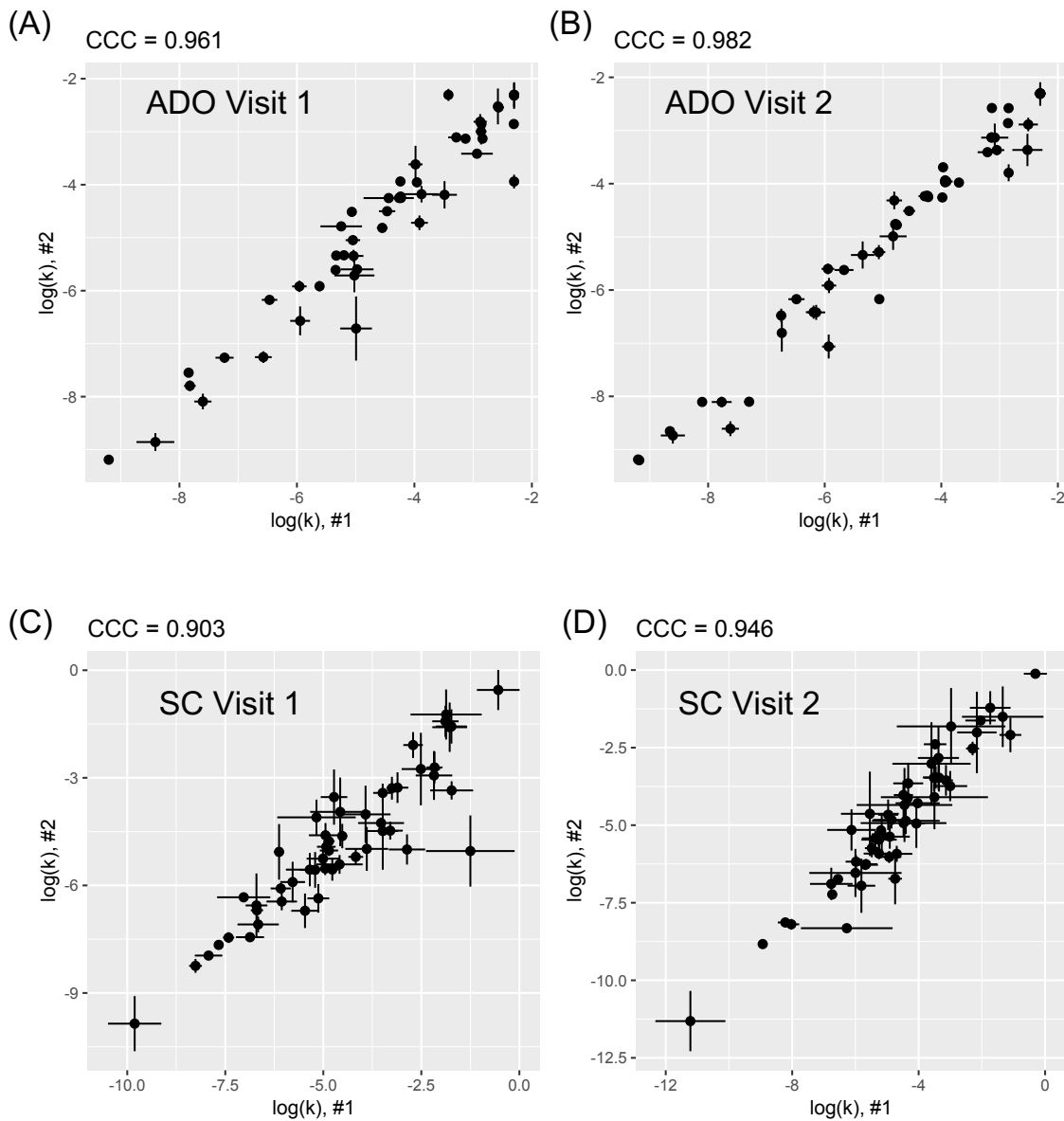
Figure S2. Within-visit test-retest reliability of temporal discounting rates (log(k), A & C) and inverse temperature parameters (β, B & D) among college students (Experiment 1) with ADO including outliers, which are indicated as red circles. See **Methods and Materials** for the description of outliers. (A & B) At visit 1, (C & D) At visit 2, which was separated by approximately one month from visit 1. Error bars on each point are Bayesian credible intervals defined in terms of ±1 standard deviation of each participant's posterior distribution of the model parameter at the end of the session. CCC: Concordance Correlation Coefficient.
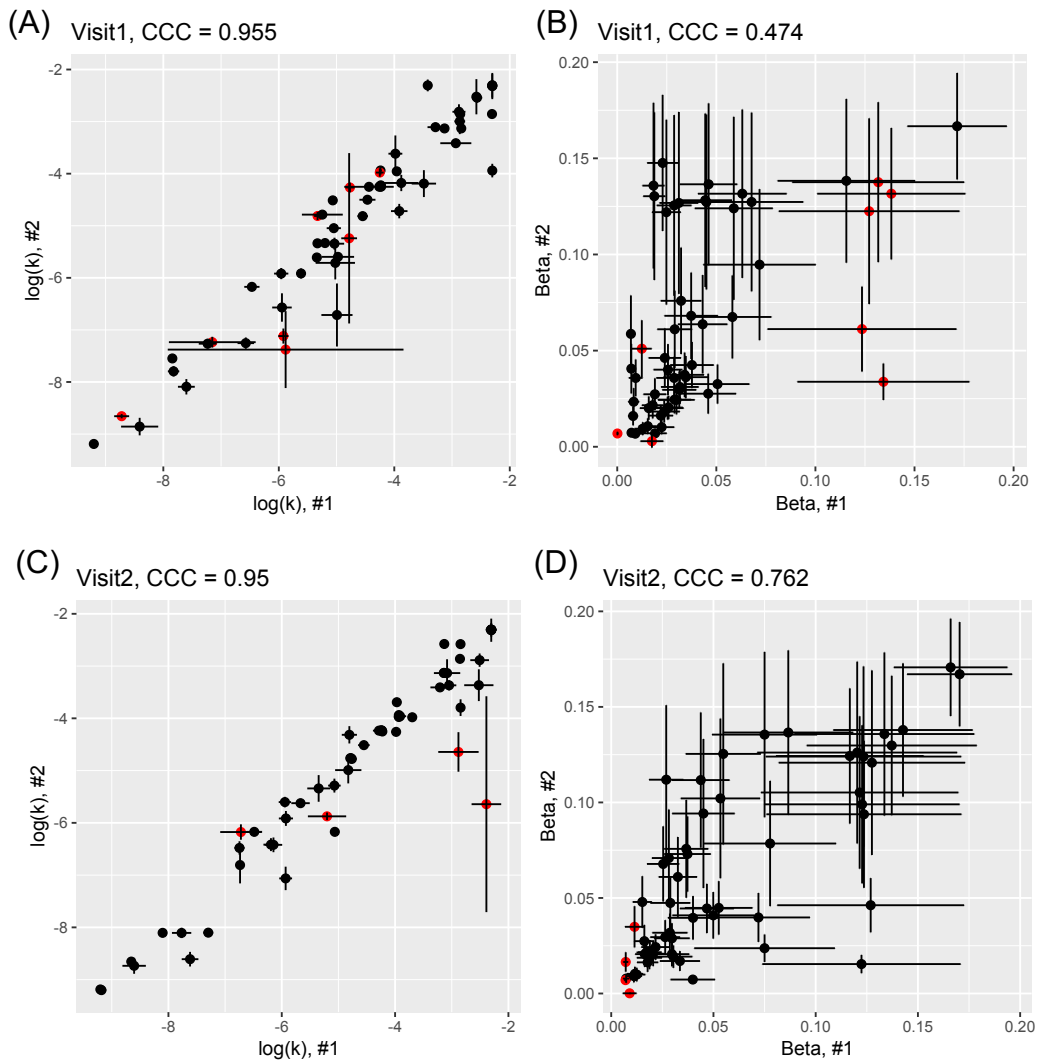
Figure S3. Within-visit test-retest reliability of temporal discounting rates (log(k), A & C) and inverse temperature parameters (β, B & D) among college students (Experiment 1) with the staircase method including outliers, which are indicated as red circles. See **Methods and Materials** for the description of outliers. (A & B) At visit 1, (C & D) At visit 2, which was separated by approximately one month from visit 1. Error bars on each point are Bayesian credible intervals defined in terms of ±1 standard deviation of each participant's posterior distribution of the model parameter at the end of the session. CCC: Concordance Correlation Coefficient.
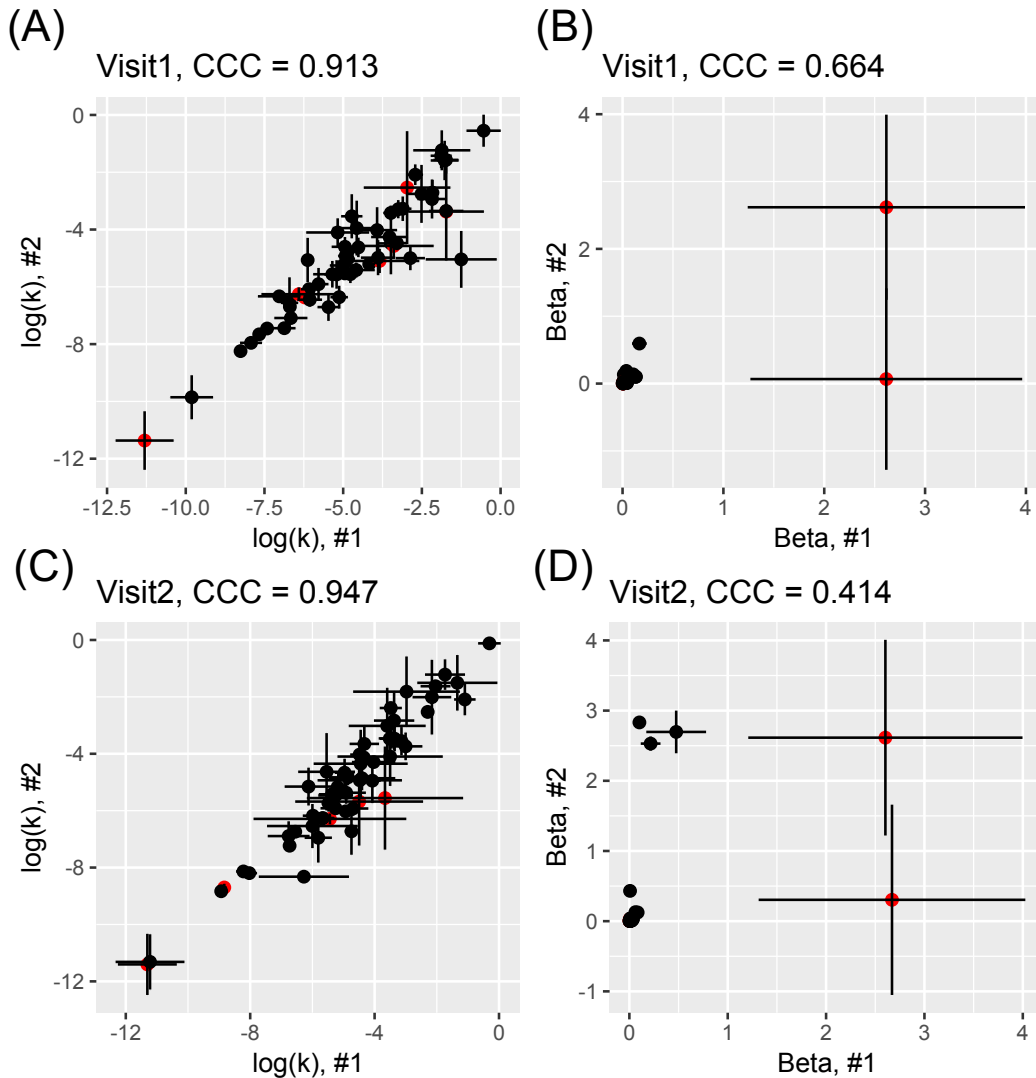
Figure S4. Comparison of ADO and staircase methods with respect to their precision of parameter estimates (Experiment 1, college students). As a measure of precision, we used the standard deviation (SD) of an individual's posterior distribution of the discounting rate parameter. Thus, the smaller SD is, the greater its precision. (A) At visit 1 (B) At visit 2, which was separated by approximately one month from visit 1. The SDs are 0.122 (visit 1) and 0.098 (visit 2) for ADO, and 0.413 (visit 1) and 0.537 (visit 2) for SC.
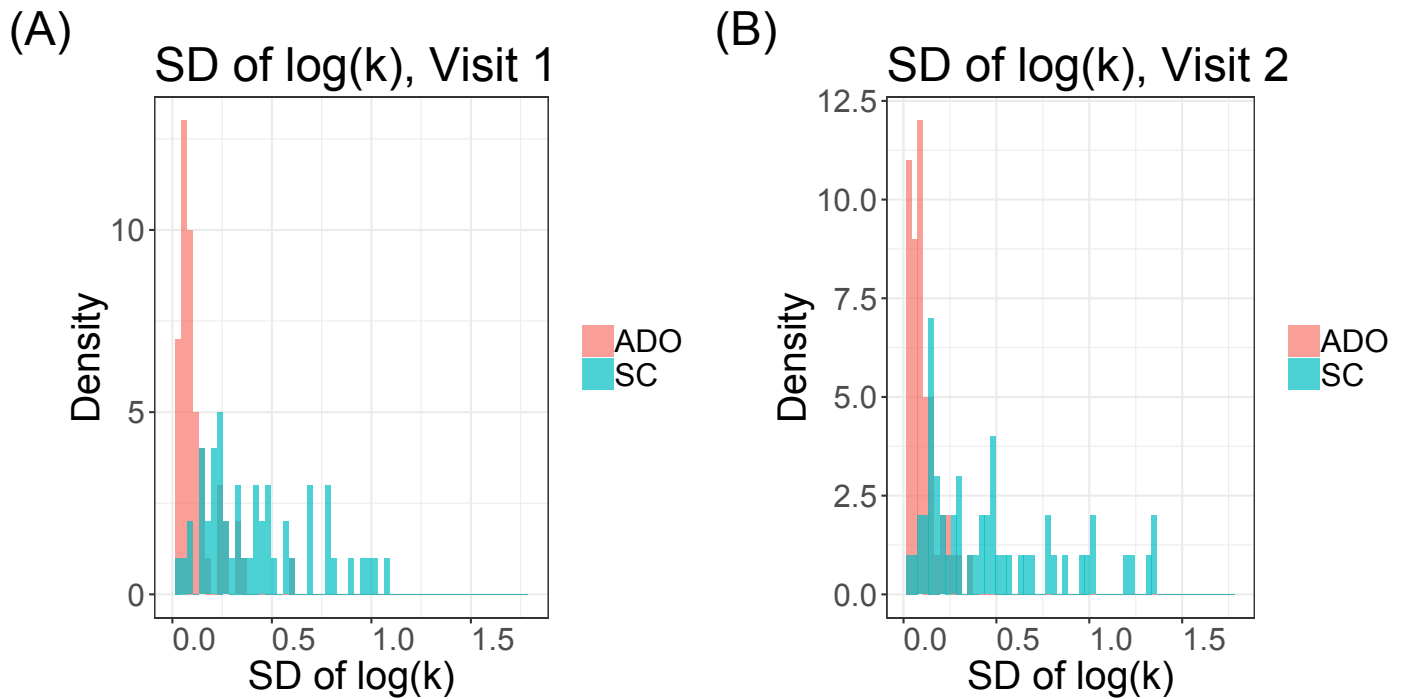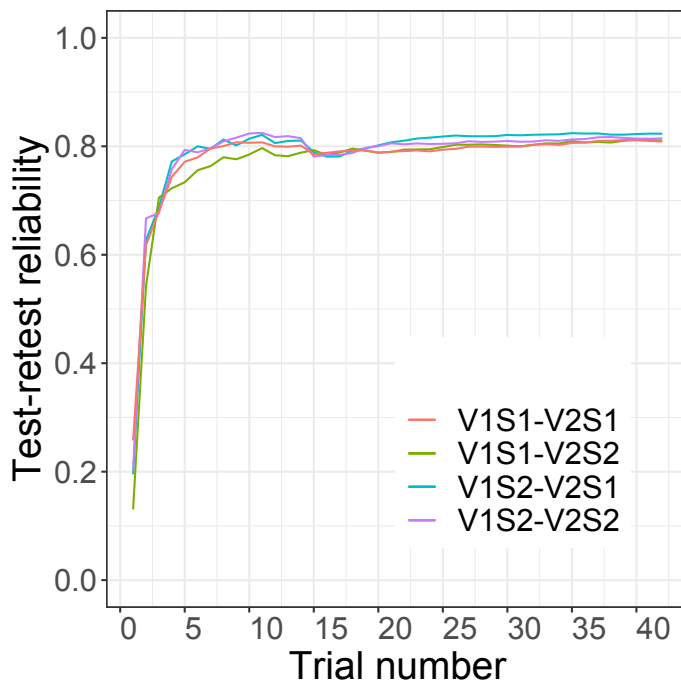
(A)

SD of log(k), Visit 1

(B)

SD of log(k), Visit 2

Figure S5. Comparison of between-visit efficiency of the ADO (A) and the staircase (B) methods (Experiment 1, college students). Between-visit efficiency is measured as the cumulative test-retest reliability in each trial (ADO) or every third trial (staircase) across two visits. The two visits were separated by approximately one month. Each line represents a different test-retest reliability comparison across the two visits (1. visit1-session1 vs visit2-session1 (red); 2. visit1-session1 vs visit2-session2 (green); 3. visit1-session2 vs visit2-session1 (cyan); 4. visit1-session2 vs visit2-session2 (purple)).
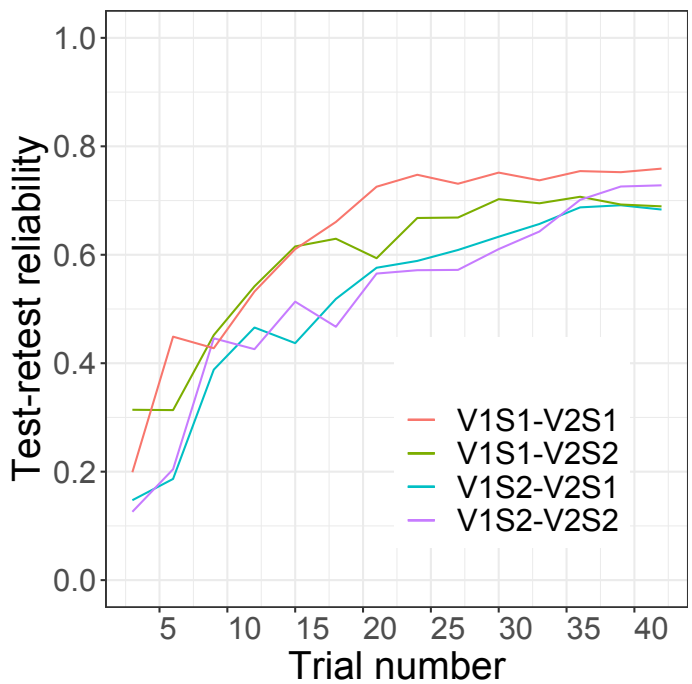
Figure S6. Test-retest reliability of discounting rates (log(k), A & C) and inverse temperature parameters (β, B & D) among patients with substance use disorders (Experiment 2) with ADO and staircase methods including outliers, which are indicated as red circles. See **Methods and Materials** for the description of outliers. (A & B) With ADO, (C & D) With the staircase method. Error bars on each point are Bayesian credible intervals defined in terms of ±1 standard deviation of each participant's posterior distribution of the model parameter at the end of the session. CCC: Concordance Correlation Coefficient.

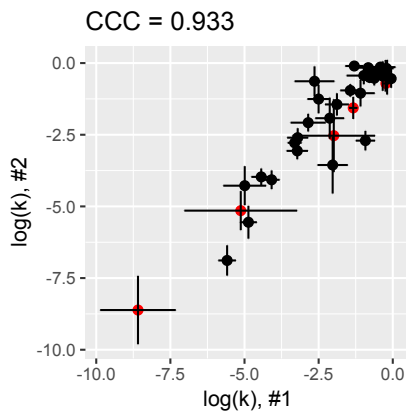# Experiment 2 (Patients with SUDs)

(A) ADO, *k*

CCC = 0.92

(B) ADO, beta

CCC = 0.688

(C) SC, *k*

CCC = 0.933
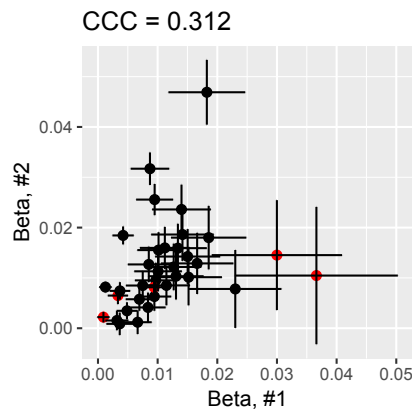
(D) SC, beta

CCC = 0.312
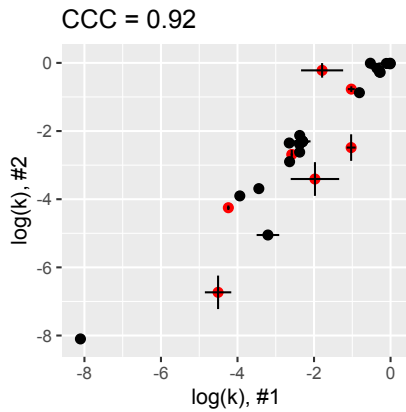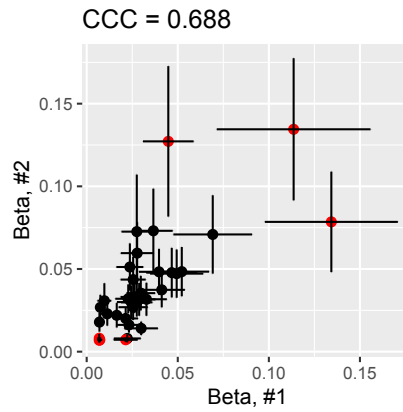
Figure S7. Reliability and efficiency of the staircase method in Experiment 2 (N=15 patients with substance use disorders (SUDs)) with ADO. Out of 35, 14 patients whose discounting rates (*k*) reached the upper bound (= 0.1) were first excluded and 6 additional patients were excluded with the 2SD rule. (A) Test-retest reliability with all 42 trials per session. Error bars on each point are Bayesian credible intervals defined in terms of ±1 standard deviation of each participant's posterior distribution of the discounting rate at the end of the session. (B) Efficiency metric using cumulative test-retest reliability across trials. Shaded regions represent the 95% frequentist confidence interval of the concordance correlation coefficient (CCC).
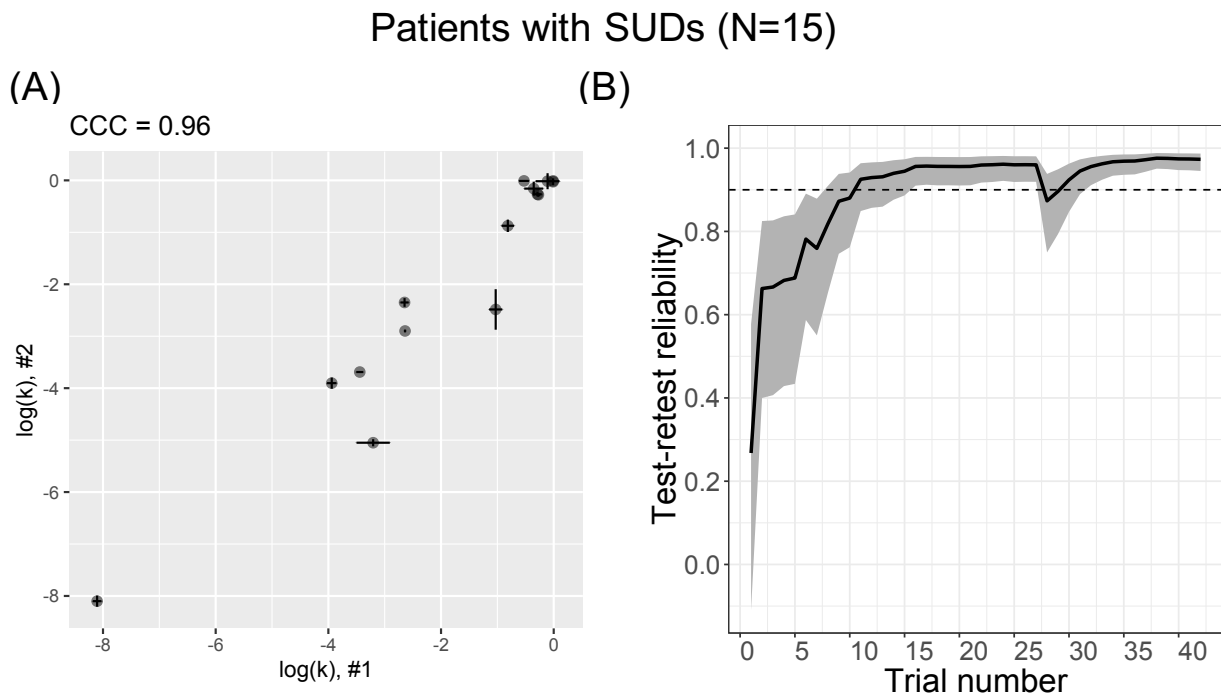

Patients with SUDs (N=15)

Figure S8. Test-retest reliability of discounting rates (log(k)) and inverse temperature parameters (β) among large online Amazon MTurk participants (Experiment 3) with ADO including outliers, which are indicated as red circles. See **Methods and Materials** for the description of outliers. Error bars on each point are Bayesian credible intervals defined in terms of ± 1 standard deviation of each participant's posterior distribution of the model parameter at the end of the session. CCC: Concordance Correlation Coefficient.
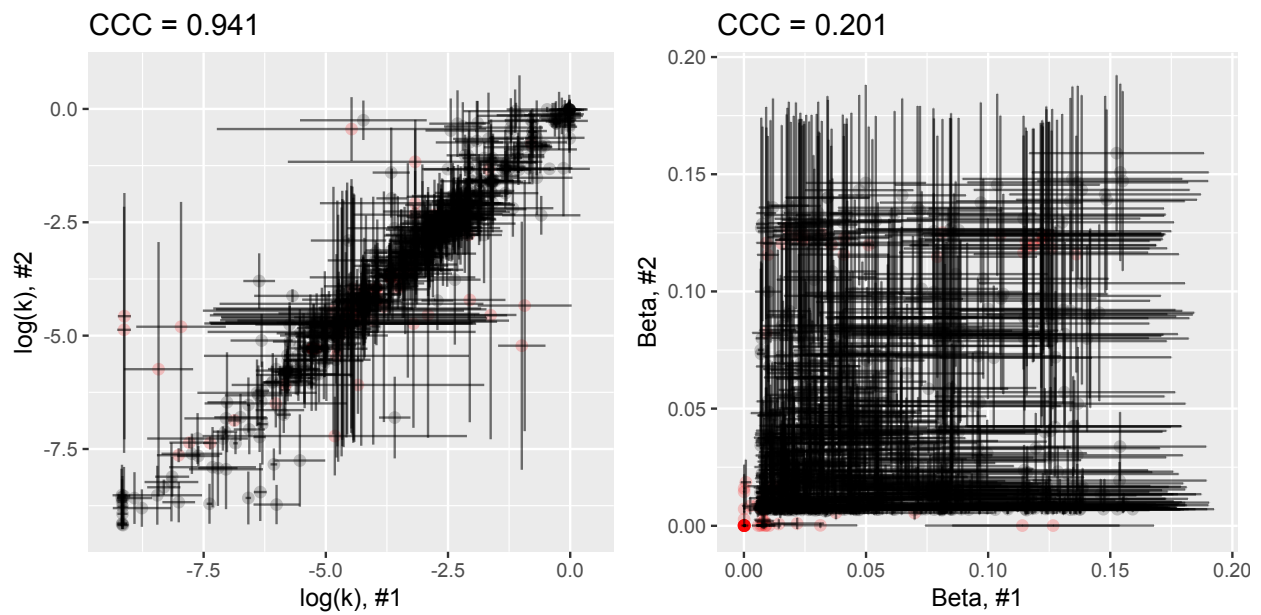
Figure S9. Trajectories of design space selection over trials with ADO and Staircase (SC) for a representative participant in Experiment 1 and visit 1 (see **Figure S10** for the trajectories of design space selection in visit 2 of the same participant). Design space refers to two experimental parameters: (1) delays for larger and later (LL) rewards and (2) amount for smaller and sooner (SS) rewards. Color bars indicate trial numbers (1 through 42 represented by a transition from blue to red).
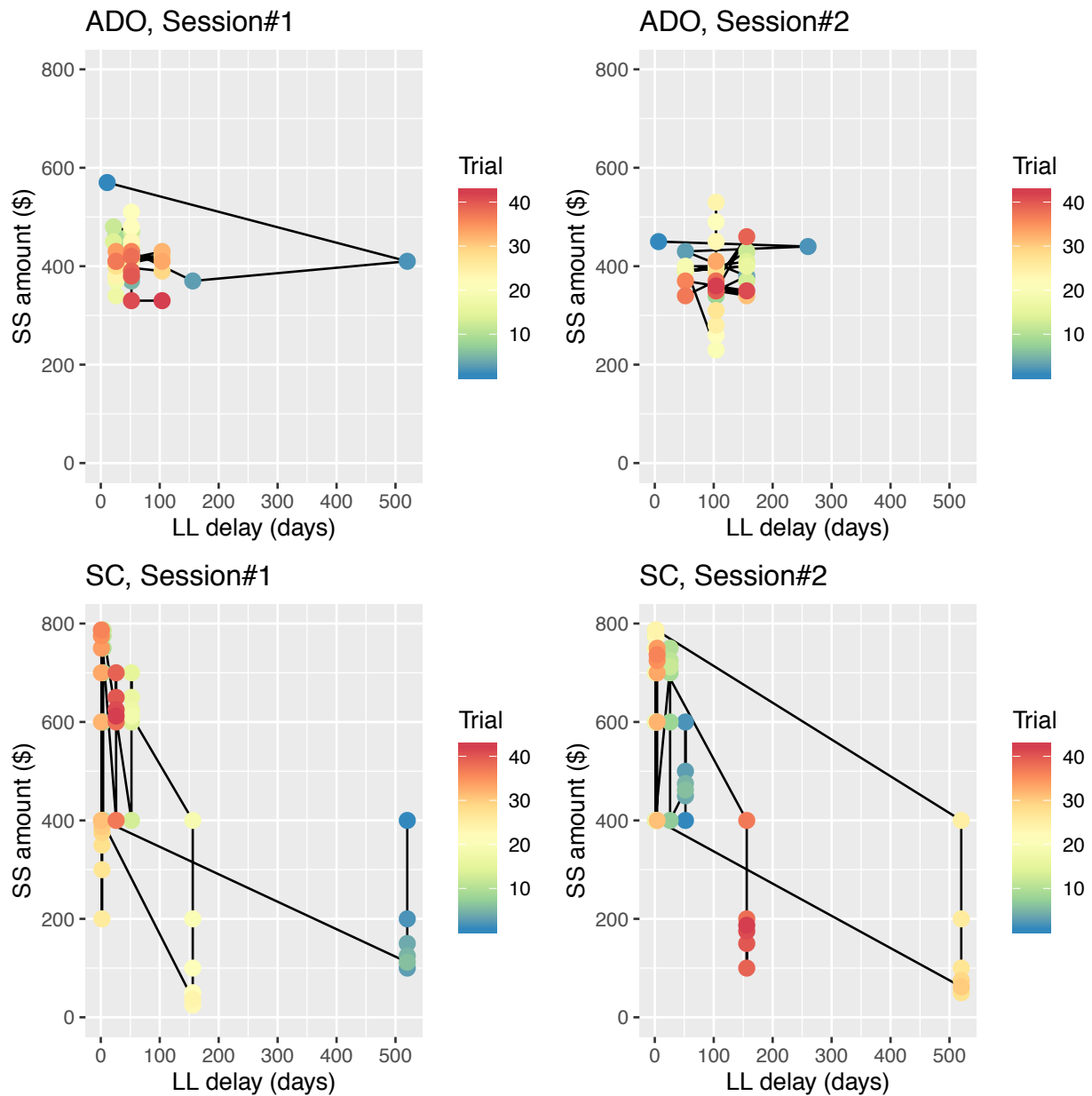
Figure S10. Trajectories of design space selection over trials with ADO and Staircase (SC) for a representative participant in Experiment 1 and visit 2 (see **Figure S9** for the trajectories of design space selection in visit 1 of the same participant). Design space refers to two experimental parameters: (1) delays for larger and later (LL) rewards and (2) amount for smaller and sooner (SS) rewards. Color bars indicate trial numbers (1 through 42 represented by transitions from blue to red).
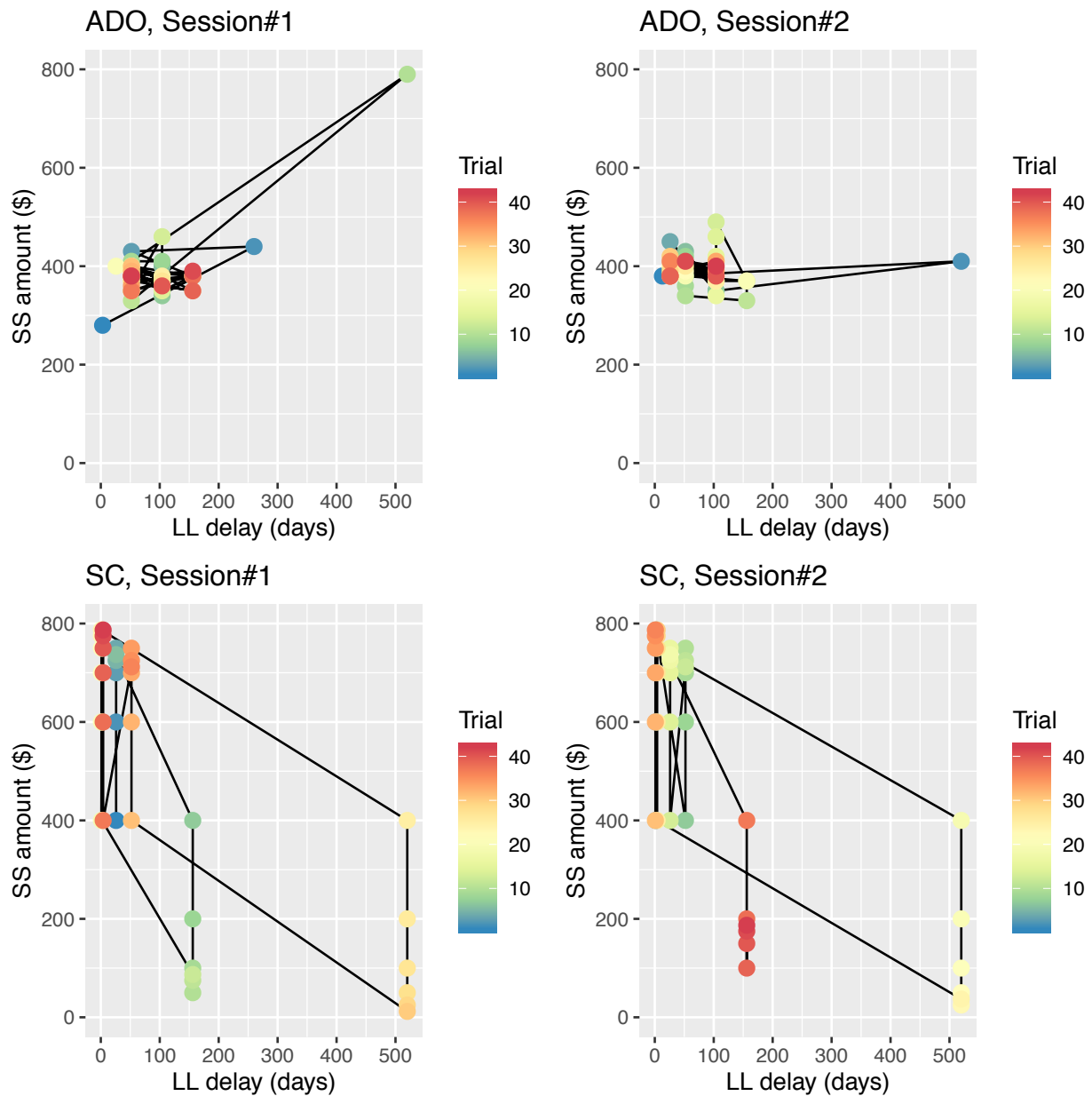
Figure S11. Analyses comparing task order regarding test-retest reliability in Experiment 1.

ADO-SC: ADO then SC, SC-ADO: SC then ADO, CCC: Concordance Correlation Coefficient.