Supplemental Material for

Shifting Negative Prospection With Online Cognitive Bias Modification:

A Randomized Controlled Trial

Jeremy W. Eberle

Charlottesville, Virginia

A predissertation research project presented to the

graduate faculty of the University of Virginia in

candidacy for the degree of Master of Arts

Department of Psychology

University of Virginia

Section S1: Method Notes

1.1 Random Assignment

We used a random number generator to assign participants to condition. Because of a programming error, about twice as many participants were assigned to Neutral Control than to other conditions before the error was detected.

1.2 Reminders

Participants had the option of receiving an email or text reminder when the next training session or the follow-up assessment was due. However, due to a programming error, for the first few months of the study some participants did not receive a reminder about the follow-up; we notified participants after we discovered the error and asked them to complete the assessment.

1.3 Partial or Multiple Completion

If participants completed only part of a training session or assessment, they resumed where they left off the next time they returned to the website. Participants who did not start the first training session by September 9, 2018, were excluded from the ITT sample, and participants who did not complete the fourth training session by this date were excluded from the PP sample. In some cases, programming errors meant that participants completed components of training sessions or assessments multiple times. We analyzed the first full set of adjacent components for a given session or assessment.

1.4 Expectancy Bias Task

The 776 college students used to determine the cutoff had completed at least one positive and one negative scenario in a 24-scenario version of the task that Namaky et al. (2019) administered in the Fall 2015 semester to the participant pool of a psychology department at a midatlantic university for course credit. The index scores based on the 24-scenario version did not significantly differ from those based on the set of 12 scenarios ultimately used in Namaky et al. In contrast to Namaky et al., to reduce response burden we included 4 of the 12 scenarios; we did not include multiple scenarios per domain or scenarios for the two domains career/work and romance/relationships (see Section S1.7 below for the scenario selection method). Moreover, to streamline the task we presented all scenarios and future events at once on a single page (rather than one scenario and future event at a time).

1.5 Internal Consistency

To compute omega total, we used 1,000, 10,000, and, if needed, 20,000 bias-corrected and accelerated bootstrap samples (until the lower and upper limits of the confidence interval for omega total were stable to two decimal places). (Note that we received warnings from the `lavaan` package, a dependency of the `MBESS` package, that the optimizer had not found a solution and that some estimated variances were negative for negative bias and positive bias in the ITT and PP samples. For the PP sample, stable confidence intervals did not emerge for

negative bias even after 40,000 bootstrap samples, and implausible estimates emerged for positive bias [i.e., 57.80, beyond omega total's range]. We also received warnings that some estimated variances were negative for self-efficacy and growth mindset in the PP sample.)

1.6 Factor Analysis

Separately for the positive events and the negative events, we conducted an exploratory factor analysis with promax rotation and maximum likelihood estimation using the `psych` package (ver. 1.8.12; Revelle, 2018) in R to determine which item had the highest loading on a single factor. We then used the `OpenMx` package to conduct a confirmatory factor analysis on the raw data with a single-factor RAM model in which we fixed the factor loading for the highest-loading item identified in the preceding exploratory analysis. To conform with the internal consistency analyses, we conducted these analyses using the complete data at baseline. During the confirmatory factor analyses, we received warnings that the Hessian did not appear to be convex, that the information matrix was not positive definite, and that standard errors for factor loadings and variances could not be computed, suggesting likely poor model fit. Good model fit is reflected, in part, by Comparative Fit Index (CFI) or Tucker Lewis Index (TLI) values of at least 0.95 or by a Root Mean Square Error of Approximation (RMSEA) value of at most 0.06 (Hu & Bentler, 1999). The highest CFI value we obtained was .04; the other CFI values were beyond the plausible range (i.e., negative). The highest TLI value we obtained was -1.89. The lowest RMSEA value we obtained was .09; one of the other RMSEA values was beyond the plausible range (i.e., greater than 1; Table S4).

1.7 Item Selection and Modification

The present study administered a subset of the self-efficacy, growth mindset, and optimism items administered by Namaky et al. (2019). Three self-efficacy items were selected from the full New General Self-Efficacy (NGSE) Scale, which contains eight 5-point Likert items ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Three growth mindset items were selected from a set of eight mindset questions adapted by Namaky et al. Two optimism items were selected from the full Life Orientation Test-Revised (LOT-R), which contains ten 5-point Likert items, six of which are scored, ranging from 0 (*strongly disagree*) to 4 (*strongly agree*).

Potential items were considered for selection based on exploratory analyses with nonmissing baseline and follow-up data from Namaky et al. (2019) for 201 college students who started the first training session in that study. The baseline and follow-up data were treated as independent observations in these analyses, which are summarized as follows. To start, a scree plot of values from principal components analyses of the full, long-format dataset was used to determine the number of components comprising each scale. The data were then split into a training set with 60% of observations and a testing set with 40% of observations, and the scree plot analysis was conducted on the training dataset, finding the same number of components. Next, the number of components was inputted into exploratory factor analyses of the training dataset to assess which items loaded onto the same number of factors. Items that did not load onto a factor at all or that had factor loadings of .3 or lower were considered for removal from the assessment battery. A full-information item factor analysis application of multidimensional item response theory was also used in the training dataset to determine which items loaded onto a single factor. Items that

did not load onto the factor at all or that had factor loadings of .3 or lower were considered for removal. Based on results from these analyses, candidate combinations of items were considered for removal, and the impact on Cronbach's alpha was assessed. The exploratory factor analyses, full-information item factor analysis, and Cronbach's alpha analyses were then conducted for the testing dataset, and an automated item selection procedure for Mokken scale analysis was done for the full dataset. Items recommended for the assessment battery based on these analyses were ultimately identified by choosing a few items from each scale (three items each for self-efficacy and optimism, four items for optimism) that maximized alpha in the full dataset. Although the initial principal components analyses in the full dataset and the Cronbach's alpha analyses in the testing, training, and full datasets implemented reverse scoring for relevant growth mindset and optimism items, the other analyses described above did not. Given this, the exploratory nature of the analyses, and lack of familiarity with some of the attempted techniques (e.g., automated item selection procedure for Mokken scale analysis), we do not claim that the recommended items are the ideal set; we simply describe the analyses and provide the data and code for transparency.

The three recommended self-efficacy items, three recommended growth mindset items, and two of the four recommended optimism items were administered in the present study. The word *that* was removed from the original NGSE item "When facing difficult tasks, I am certain that I will accomplish them." In addition, the word *with* was replaced with *for* in the original LOT-R item "If something can go wrong for me, it will."

The present study also administered a subset of the Expectancy Bias Task scenarios administered by Namaky et al. (2019). Four out of the 12 scenarios used in Namaky et al. were selected for the present study by excluding scenarios whose items contributed least to Cronbach's alpha until the remaining items retained at least acceptable internal consistency. In addition, it was determined that analyzing only the items from these four scenarios yielded similar expectancy bias outcomes in the data from Namaky et al. as analyzing the items from all 12 scenarios. The expectancy bias data from Namaky et al. and the code used to select these four scenarios is not available.

1.8 Missing Letters

In CBM tasks with two letters missing (Sessions 3-4), any of the letters could be missing; no logic prevented the missing letters from being consecutive letters in the fragment.

1.9 Comprehension Questions

The mean proportion of scenarios with comprehension questions was two thirds. A comprehension question was presented after a scenario if a random number between 0 and 1 generated for the scenario exceeded .33. This resulted in a distribution of the proportion of scenarios followed by comprehension questions centered around two thirds. Inserting comprehension questions for only two thirds of scenarios reduced the total training time while still encouraging participants to attend to the scenarios' resolutions.

1.10 Baseline Demographic Differences

Using the `stats` package in R (ver. 3.5.3), for age (computed from reported birth year and

timestamped year) we conducted a one-way analysis of variance and, because the normality assumption was violated, a Kruskal-Wallis rank-sum test. We conducted Fisher's exact tests for gender, race, country, and marital status; chi-square tests of independence for ethnicity and employment; and Kruskal-Wallis rank-sum tests for education and income. Although these significance tests are illogical (any significant findings would be Type I errors by definition given random assignment to condition), for transparency we report their results below.

No significant findings emerged. The one-way analysis of variance and Kruskal-Wallis rank-sum test indicated that conditions did not significantly differ by age: $F(4, 951) = 0.24$, $p = .916$; $\chi^2(4) = 0.99$, $p = .911$. The Fisher's exact tests indicated that they did not significantly differ by gender, $p = .877$, race, $p = .225$, country, $p = .741$, or marital status, $p = .691$. The chi-square tests of independence indicated that they did not significantly differ by ethnicity, $\chi^2(4, N = 893) = 3.16$, $p = .532$, or employment, $\chi^2(28, N = 946) = 21.99$, $p = .782$. Finally, the Kruskal-Wallis rank-sum tests revealed that they did not significantly differ by education, $\chi^2(4) = 1.72$, $p = .787$, or income, $\chi^2(4) = 0.98$, $p = .913$.

1.11 Auxiliary Variables

Specifically, we used the `DescTools` package (ver. 0.99.28; Signorell, 2019) in R (ver. 3.5.1.; R Core Team, 2018) to compute Goodman and Kruskal's gamma for age, which indicated a negative ordinal association such that participants ranking higher in age ranked significantly lower in their number of missing sessions, $G = -0.06$, 95% CI [-0.11, -0.01], and to conduct a Jonckheere-Terpstra test for level of education, which indicated an ordering of medians such that the median number of missing sessions significantly decreased across higher levels of education, $J = 140{,}980$, $p = .002$. A Jonckheere-Terpstra test for income was nonsignificant, $J = 142{,}210$, $p = .276$, as were Wilcoxon rank-sum tests for gender and ethnicity and Kruskal-Wallis rank-sum tests for race, marital status, employment, and country run with the `stats` package in R: $W = 89{,}344$, $p = .351$; $W = 20{,}222$, $p = .839$; $\chi^2(4) = 5.15$, $p = .272$; $\chi^2(9) = 7.51$, $p = .584$; $\chi^2(7) = 9.16$, $p = .242$; $\chi^2(8) = 10.30$, $p = .245$.

1.12 Multiple Imputation

We saved the first dataset at the 10,000th iteration and additional datasets every 250th iteration thereafter. The potential scale reduction factor (< 1.05), trace plot, and autocorrelation plot for all imputation model parameters were examined to ensure that they converged and that the imputed datasets were selected independently from the missing data's posterior predictive distribution (following the steps in Grund et al., 2016). Although the imputation model generated some values beyond the possible scale ranges, the mean percentage of out-of-range values per imputed dataset did not exceed about 15% for any scale (Table S5, Figures S1-S7); we considered this rate small enough to only trivially inflate variance estimates (Enders, 2010, p. 265).

1.13 Iatrogenic Effects

To translate index scores into the positive range, we added 8 to each score such that the new possible range was from 2 to 14. (Although we could have added 7 to obtain a range from 1 to 13, it was thought during software development that the index scores could range from -7 to 7

such that adding 8 was required.)

1.14 Exploratory Analyses of Anxiety and Depression Composite

Using the imputed datasets we had already generated for our other longitudinal analyses based on a multiple imputation model in which anxiety and depression were included as separate Level 1 outcomes, we summed the anxiety and depression scores in each dataset to form an anxiety and depression composite, representing the four-item PHQ-4 total score. Although to our knowledge a bifactor model (see Reise, Moore, & Haviland, 2010) has not been fit to the PHQ-4, its total score is thought to be a composite measure of anxiety and depression symptoms (Kroenke, Baye, & Lourens, 2019), and authors who have fit unidimensional and two-dimensional models of the PHQ-4 claim that the total score can be used in addition to the anxiety and depression subscales (Löwe et al., 2010). Moreover, a bifactor model fit to the sum of the PHQ-9 and GAD-7 found sufficient unidimensionality to create the PHQ-Anxiety and Depression Scale (Kroenke et al., 2016). After computing the anxiety and depression composite, we analyzed it as a longitudinal outcome using the same methods described in the Multilevel Modeling and Effect Size sections.

These analyses of the anxiety and depression composite revealed the same pattern of results as the analyses that treated anxiety and depression as separate outcomes. Effect sizes for the simple effects of time in each condition in the analyses of the composite are greater than or fall between those in the separate anxiety and depression analyses (Table S8). Effect sizes for the Condition × Time interactions fall between those in the separate anxiety and depression analyses (Table S9). Because the analyses of the composite are exploratory, we do not report *p*-values for them.

Section S2: Preregistration Deviations

2.1 Comparing Positive Conditions as Nondirectional Test

Our preregistration (osf.io/jrst6) states that we will test for differential changes in outcomes between the two positive conditions using the same approach as our comparisons of other conditions. We called the comparison of the positive conditions, for which we did not have a directional hypothesis, an *exploratory* analysis and the comparisons of the other conditions, for which we had directional hypotheses, *confirmatory* analyses. However, all of these tests are confirmatory. The former test is simply nondirectional, whereas the latter tests are directional.

2.2 Analyzing Sums for Anxiety and Depression

Our preregistered analysis plan states that we will analyze the means rather than the sums of the anxiety and depression items. We analyzed the sums to follow the original scoring procedures for the GAD-2 and PHQ-2, which call for sums. Moreover, there are no item-level missing data for these measures; thus, the sums are interpretable.

2.3 Not Conducting Mediation Analyses

Our preregistered analysis plan states that we will use the `mediation` package in R following the method described by Tingley, Yamamoto, Hirose, Keele, and Imai (2014) to assess for mediation. These analyses have not been conducted yet and we elected not to include them here in part due to space concerns and in part because we expect that other approaches we have since identified (e.g., parallel process growth curve modeling, cross-lagged panel modeling) may more appropriately assess our mediation hypotheses.

2.4 Not Removing Outliers or Transforming Outcomes

Our preregistered analysis plan states that for each dependent variable we will remove outliers that exceed 3 median absolute deviations from the median and consider transforming the variable if model diagnostics indicate an assumption violation. Given that age and education were related to monotone, scale-level missingness (i.e., scale scores were not missing completely at random), we did not expect univariate normality for each scale prior to imputation, so we did not test for outliers or transform scale scores as planned. Nor did we believe that it would be appropriate after imputation to exclude outliers or transform imputed values. We assumed multivariate normal and missing-at-random data for the imputation and subsequent analysis models, but for transparency we show the univariate distributions of each scale at each time point before and after imputation in Figures S1-S7. We also show for each scale at each time point the mean percentage of out-of-range imputed values per dataset, the mean minimum/maximum imputed values per dataset, and the absolute minimum/maximum values across datasets in Table S3.

2.5 Analyzing Follow-Up Data

Our preregistered analysis plan does not state that we will analyze the follow-up time point because during preregistration we believed there were too few follow-up data to analyze.

However, we have since learned that multiple imputation and maximum likelihood can handle high rates of missing data (e.g., even 80%) provided that their assumed missing data mechanism (missing at random) is satisfied.

2.6 Using Restricted Maximum Likelihood Estimation With `nlme`

Our preregistered analysis plan states that we will use the `lme4` package (Bates, Maechler, Bolker, & Walker, 2015) in R to fit mixed-effects models with maximum likelihood estimation. The `nlme` package also fits these models. Maximum likelihood yields biased variances and covariances, whereas restricted maximum likelihood does not (Enders, 2010, p. 80).

2.7 Assuming Linear Trajectories

Our preregistered analysis plan states that we will compare linear, quadratic, cubic, and quartic models to determine the shape that best describes the trajectory for each outcome over time. We decided to include only the linear term for time because our hypotheses concerned differential rates of change (velocity) between conditions, not differential acceleration, jerk, or jounce, which would require the higher-order terms.

2.8 Including Condition × Time Interaction

Our preregistered analysis plan states that we will assess the fixed effect of condition on change scores from one session to the next while controlling for the fixed effect of time. Because we were interested in differences between conditions on the overall trajectory of change rather than on session-by-session change, we analyzed scores at each time point rather than computing change scores, included the Condition × Time interaction, and interpreted this interaction rather than the fixed effect of condition.

2.9 Analyzing Anxiety and Depression Composite

Our preregistered analysis plan states that we will analyze anxiety and depression symptoms as separate outcomes. After we conducted these confirmatory analyses, we learned that Namaky et al. (2019) found in exploratory analyses that the positive conditions improved on an anxiety and depression composite score, the 14-item sum of the Anxiety and Depression subscales of the Depression Anxiety Stress Scales-21, significantly more than the neutral condition (N. Namaky, personal communication, September 1, 2019). To explore this possibility in the present study, we also analyzed an anxiety and depression composite, the four-item sum of the PHQ-4.

Section S3: Guide to Open Data and Open Materials

3.1 Open Data

Raw data received from the website server are stored in an SQL database or in JSON files. The present manuscript uses data in CSV format that MindTrails project coordinators exported from these sources using queries on the date indicated in the file name of each CSV raw data file (see below). The CSV files retain variable names and values found in the SQL database and JSON files with the exception of identifiable data about participants, which the project coordinators removed, misaligned columns corrected in one table, and two renamed variables in one table (these exceptions are described in the syntax file 1_get_raw_data.rtf—see below).

Data in CSV format and R code used in this manuscript are available at https://osf.io/jp5ws/. Raw data are in the folder Data/Raw, intermediate data are in the folder Data/Temp, clean data are in the folder Data/Clean, and imputed data (R list objects) are in the folder Data/Imputed. The folders Data/Raw and Data/Clean contain codebooks for the data tables therein.

R code used to transform raw data into the other forms and to conduct analyses is in the folder Syntax. Scripts should be consulted in the numbered order, starting with 0_item_selection.R, which shows analyses used to select items from Namaky et al. (2019) for the present study. The script 1_get_raw_data.R describes how the files containing raw data collected from the present study were obtained, and the subsequent files show the cleaning of these data and analyses of them. See Table S10 for manuscript content indexed by R script and code section.

Other files exported by the R scripts are in the folder Results. Table S1 in CSV format is in the subfolder Table_S1. Histograms in PDF format used for Figures S1-7; results in TXT format used for Table S5; and imputation diagnostics (potential scale reduction factors in TXT format, trace and autocorrelation plots in PDF format) are in the subfolder Imputation. Results in TXT format used for Tables 1, 2, S6, and S7 are in the subfolder Longitudinal_Outcome.

3.2 Open Materials

The website code used to run the Future Thinking study is available in the core and templeton folders of the MindTrails Project repository on GitHub. The latest release of MindTrails code is available at https://github.com/TeachmanLab/MindTrails; the code at the time the present study launched is available at https://github.com/TeachmanLab/MindTrails/tree/templeton_release.

The scenarios and comprehension questions, available in CSV format in the templeton_release repository above, can also be found at https://osf.io/jp5ws/ in the folder Materials/Scenarios. For the Positive Prospection, Positive Prospection + Negation, 50/50 Blocked, and 50/50 Random conditions, see scenarios.csv. For the Neutral Control condition, see scenarios_neutral.csv.

Screenshots of the study website in PDF format can be found at https://osf.io/jp5ws/ in the folder Materials/Screenshots.

Table S1

Raw Means and Standard Deviations of Outcomes by Condition Over Time for the Intent-To-Treat Sample

Outcome	Assessment	Pos. + Negation			Positive			50/50 Blocked			50/50 Random			Neutral Control		
		n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Positive Bias	Baseline[a]	147	3.34	0.98	177	3.42	0.95	146	3.44	0.87	173	3.24	0.98	315	3.27	1.00
	Session 1	114	4.27	1.09	142	4.48	1.04	113	4.19	0.97	139	4.09	1.08	274	3.84	1.12
	Session 2	66	4.93	1.12	79	5.00	0.97	65	4.37	1.13	82	4.54	0.98	172	4.20	1.13
	Session 3	43	4.97	1.19	59	5.36	0.94	53	4.38	1.26	66	4.81	1.04	137	4.41	1.18
	Session 4	36	5.06	1.22	45	5.28	0.91	42	4.70	1.15	50	4.76	1.24	116	4.41	1.15
	Follow-up	13	4.96	1.25	23	4.91	1.00	25	4.73	1.15	28	4.75	1.20	60	4.38	1.37
Negative Bias	Baseline[a]	147	4.21	1.02	177	4.12	0.90	146	4.17	1.08	173	4.13	0.91	315	4.14	0.95
	Session 1	114	3.58	1.11	142	3.30	0.97	113	3.51	1.09	139	3.70	0.99	274	3.73	1.04
	Session 2	66	3.06	1.11	79	2.88	0.94	65	3.69	1.12	82	3.50	0.99	172	3.57	1.09
	Session 3	43	3.11	1.18	59	2.81	0.88	53	3.58	1.10	66	3.38	1.03	137	3.43	1.08
	Session 4	36	3.06	1.08	45	2.75	1.00	42	3.25	1.14	50	3.39	1.21	116	3.34	1.12
	Follow-up	13	3.38	1.15	23	2.86	1.06	25	3.19	1.07	28	3.17	1.03	60	3.26	1.09
Anxiety	Baseline[b]	147	3.79	1.78	177	3.49	1.93	146	3.29	1.74	173	3.20	1.96	315	3.62	1.89
	Session 2	66	2.77	1.67	79	2.84	1.85	65	3.00	1.88	82	2.96	1.64	172	2.98	1.76
	Session 4	36	2.53	1.66	45	2.67	1.89	42	2.24	1.72	50	2.82	1.72	116	2.67	1.75
	Follow-up	13	2.92	2.18	23	2.39	1.80	25	2.56	1.89	28	2.25	1.71	60	2.68	1.90
Depression	Baseline[b]	147	2.71	1.92	177	2.76	1.93	146	2.58	1.74	173	3.03	2.00	315	2.95	1.93
	Session 2	66	2.02	1.63	79	2.11	1.70	65	2.46	1.79	82	2.38	1.70	172	2.19	1.80
	Session 4	36	1.94	2.03	45	1.71	1.83	42	1.90	2.06	50	2.14	1.83	116	2.09	1.85
	Follow-up	13	2.38	2.26	23	1.74	2.00	25	1.80	1.76	28	2.50	1.86	60	2.23	1.99
Self-Efficacy	Baseline[b]	147	2.31	0.85	176[d]	2.23	0.89	146	2.28	0.92	173	2.14	0.95	314[c]	2.21	0.91
	Session 2	66	2.77	0.89	79	2.91	0.65	65	2.49	0.88	82	2.65	0.84	172	2.55	0.87
	Session 4	36	2.76	0.94	45	2.97	0.59	42	2.67	0.84	50	2.75	0.83	116	2.64	0.85
	Follow-up	13	2.33	1.21	23	2.96	0.71	25	2.84	0.62	28	2.51	0.99	60	2.63	0.98
Growth Mindset	Baseline[b]	147	2.06	0.83	176[d]	2.12	0.90	146	2.16	0.81	173	2.09	0.95	315	2.00	0.88
	Session 2	66	2.55	0.85	79	2.43	0.92	65	2.50	0.79	82	2.59	0.92	172	2.28	0.89
	Session 4	36	2.42	0.92	45	2.52	1.01	42	2.56	0.78	50	2.71	0.96	116	2.29	0.95
	Follow-up	13	2.44	0.98	23	2.57	0.76	25	2.76	0.78	28	2.71	1.04	60	2.28	0.94

Optimism	Baseline[b]	147	1.59	0.95	176[d]	1.68	0.89	146	1.71	0.97	173	1.66	0.97	315	1.52	0.97
	Session 2	66	2.12	0.93	79	2.16	0.92	65	1.97	0.97	82	1.94	0.98	172	1.86	1.03
	Session 4	36	2.26	0.91	45	2.33	0.90	42	2.23	0.93	50	1.93	1.05	116	1.99	0.97
	Follow-up	13	2.19	1.03	23	2.57	0.84	25	2.48	0.85	28	2.23	0.89	60	1.92	1.16

[a]Screening.
[b]Pretreatment.
[c]One participant refused to answer all self-efficacy items at baseline.
[d]One participant refused to answer all self-efficacy, growth mindset, and optimism items at baseline.

Table S2

Demographic Characteristics by Condition for the Intent-To-Treat Sample

Characteristic	Pos. + Negation (n = 147)	Positive (n = 180)	50/50 Blocked (n = 146)	50/50 Random (n = 173)	Neutral Control (n = 315)
Age (years): *M* (*SD*)	41.32 (13.75)[b]	40.66 (13.05)[a]	40.12 (13.45)	41.39 (13.78)	41.06 (13.28)[cd]
Gender: %					
Female	73.5	74.6	71.2	75.1	72.4
Male	26.5	23.7	28.8	23.7	25.7
Transgender	0.0	0.0	0.0	0.6	0.0
Other	0.0	0.0	0.0	0.0	0.6
Unknown	0.0	0.6	0.0	0.0	0.0
Prefer not to answer	0.0	1.1	0.0	0.6	1.3
Race: %					
American Indian/Alaska Native	0.0	0.0	0.7	0.0	0.0
Black/African origin	4.1	1.7	3.4	2.3	2.2
East Asian	0.7	4.5	4.1	2.3	1.6
Native Hawaiian/Pacific Islander	0.0	0.0	0.0	0.0	0.3
South Asian	0.7	4.5	0.7	1.2	2.2
White/European origin	88.4	84.2	85.6	86.7	87.3
Other or unknown	3.4	1.7	4.1	4.0	4.4
Prefer not to answer	2.7	3.4	1.4	3.5	1.9
Ethnicity: %					
Hispanic or Latino	2.7	4.0	6.8	5.2	5.4
Not Hispanic or Latino	90.5	89.3	87.7	87.9	87.3
Unknown	1.4	3.4	4.1	2.9	2.9
Prefer not to answer	5.4	3.4	1.4	4.0	4.4
Country: %					
United States	81.0	82.5	82.2	83.8	85.4
United Kingdom	4.1	4.5	2.7	2.3	3.2
Canada	0.7	3.4	2.1	2.9	1.6
Australia	4.1	1.1	0.7	2.3	1.6
Russian Federation	0.7	0.6	0.7	1.2	1.0

Other[e]	8.8	6.2	11.6	6.9	7.0
Prefer not to answer	0.7	1.7	0.0	0.6	0.3
Education: %					
No high school	0.0	0.0	0.0	0.0	0.0
Some high school	0.0	1.7	1.4	0.6	2.2
High school graduate	3.4	3.4	5.5	4.6	4.4
Some college	18.4	20.9	17.1	18.5	21.0
College graduate	25.9	26.6	28.1	32.9	25.1
Some graduate school	9.5	5.1	6.2	3.5	8.3
Advanced degree	41.5	40.7	40.4	38.7	38.1
Unknown	0.0	0.6	0.0	0.0	0.0
Prefer not to answer	1.4	1.1	1.4	1.2	1.0
Employment status: %					
Student	13.6	14.7	8.2	11.6	9.5
Homemaker	5.4	4.0	6.8	7.5	4.8
Unemployed	8.8	4.5	11.6	10.4	9.5
Working part time	12.2	13.0	10.3	12.1	10.8
Working full time	52.4	55.9	51.4	49.1	55.6
Retired	3.4	3.4	5.5	4.6	4.8
Other	3.4	1.7	4.8	4.0	4.1
Unknown	0.0	0.6	0.0	0.0	0.0
Prefer not to answer	0.7	2.3	1.4	0.6	1.0
Annual Income: %					
Less than $5,000	6.8	2.8	4.1	4.0	3.2
$5,000-$11,999	2.0	2.3	4.8	2.9	3.2
$12,000-$15,999	0.7	1.1	0.7	2.9	3.5
$16,000-$24,999	3.4	5.6	6.8	2.9	4.4
$25,000-$34,999	8.8	5.6	6.8	5.8	6.0
$35,000-$49,999	9.5	10.7	8.2	10.4	5.4
$50,000-$74,999	13.6	15.3	11.0	14.5	15.6
$75,000-$99,999	9.5	10.2	10.3	11.6	10.2
$100,000-$149,999	14.3	18.1	16.4	16.2	14.9
$150,000-$199,999	6.8	4.5	6.8	5.2	7.0

$200,000-$249,999	2.7	3.4	3.4	1.2	4.8
$250,000 or greater	5.4	4.5	6.8	9.2	5.7
Unknown	4.8	5.1	4.8	4.0	2.9
Prefer not to answer	11.6	10.7	8.9	9.2	13.3
Marital Status					
Single	21.8	20.3	26.7	23.7	23.5
Dating	8.2	10.2	8.2	6.4	7.9
Engaged	1.4	2.8	2.7	4.6	2.2
In a marriage-like relationship	7.5	11.3	8.9	5.2	6.3
Married	42.2	42.9	37.7	46.8	47.3
In a domestic/civil union	1.4	0.0	1.4	2.3	1.0
Separated	3.4	1.1	1.4	1.7	2.5
Divorced	11.6	6.2	8.2	6.4	7.9
Widow/widower	0.7	0.6	2.1	0.6	0.3
Other	0.7	1.7	1.4	1.2	0.6
Unknown	0.0	0.6	0.0	0.0	0.0
Prefer not to answer	1.4	2.3	1.4	1.2	0.3

[a]Including two participants who reported a birth year of 2000, suggesting an age of 17.
[b]Excluding one participant who reported a birth year of 1900. The participant is included in all other analyses.
[c]Excluding one participant who reported a birth year of 2017. The participant is included in all other analyses.
[d]Including one participant who reported a birth year of 2002, suggesting an age of 15.
[e]Germany (*n* = 6), Ireland (*n* = 5), New Zealand (*n* = 5), South Africa (*n* = 5), Netherlands (*n* = 4), Switzerland (*n* = 4), Brazil (*n* = 3), Colombia (*n* = 3), Greece (*n* = 3), Italy (*n* = 3), Romania (*n* = 3), Sweden (*n* = 3), Croatia (*n* = 2), Czech Republic (*n* = 2), Denmark (*n* = 2), France (*n* = 2), Mexico (*n* = 2), Spain (*n* = 2), Antarctica (*n* = 1), Ecuador (*n* = 1), Estonia (*n* = 1), Gibraltar (*n* = 1), Guyana (*n* = 1), India (*n* = 1), Japan (*n* = 1), Kuwait (*n* = 1), Malaysia (*n* = 1), Republic of Moldova (*n* = 1), Norway (*n* = 1), Panama (*n* = 1), Puerto Rico (*n* = 1), Slovakia (*n* = 1), Turkey (*n* = 1), Ukraine (*n* = 1).

Table S3

Standardized Cronbach's Alphas at Baseline for the Intent-To-Treat and Per-Protocol Samples

Outcome	Sample	Estimate
Positive Bias	ITT	.31
	PP	.13
Negative Bias	ITT	.24
	PP	.19
Depression	ITT	.81
	PP	.83
Anxiety	ITT	.82
	PP	.83
Self-Efficacy	ITT	.81
	PP	.82
Growth Mindset	ITT	.81
	PP	.83
Optimism	ITT	.80
	PP	.81

Note. ITT = intent-to-treat; PP = per-protocol.

Table S4

Fit Indices for Confirmatory Factor Analyses at Baseline for the Intent-To-Treat and Per-Protocol Samples

Outcome	Sample	χ^2	χ^2/df	df	*p*	CFI	TLI	RMSEA, 95% CI
Positive Bias	ITT	55.17	27.59	2	< .001	0.04	-1.89	0.17 [0.12, 0.21]
	PP	6.30	3.15	2	.043	-15.31[a]	-47.93	0.09 [0.00, 0.18]
Negative Bias	ITT	6748.28	3374.14	2	< .001	-129.71[a]	-391.14	1.88 [1.84, 1.93]
	PP	12.94	6.47	2	.002	-0.34[a]	-3.01	0.14 [0.06, 0.23]

Note. ITT = intent-to-treat; PP = per-protocol; CFI = Comparative Fit Index; TLI = Tucker Lewis Index; RMSEA = Root Mean Square Error of Approximation.

[a]Beyond 0-1 range of CFI (Hu & Bentler, 1999).

Table S5

Extreme Scores Across 100 Imputations of Datasets for the Intent-To-Treat Sample

Outcome	Assessment	Plausible Range	*M* % Below	Minimum Score		*M* % Above	Maximum Score	
				M	Absolute		*M*	Absolute
Combined-Level Dataset								
Positive Bias	Baseline[a]	[1, 7]	0.00	1.00	1.00	0.00	6.00	6.00
	Session 1		0.05	0.87	-0.21	0.02	7.05	7.82
	Session 2		0.09	0.76	-1.29	0.28	7.46	8.79
	Session 3		0.09	0.83	-1.08	1.70	8.33	9.70
	Session 4		0.12	0.64	-1.38	5.57	9.24	10.95
	Follow-up		0.17	0.52	-1.41	13.54	10.12	11.80
Negative Bias	Baseline[a]	[1, 7]	0.00	2.00	2.00	0.00	7.00	7.00
	Session 1		0.03	0.91	0.10	0.02	7.02	7.49
	Session 2		0.32	0.46	-0.49	0.03	6.79	8.02
	Session 3		1.14	-0.06	-1.27	0.03	6.92	8.34
	Session 4		3.13	-0.66	-2.14	0.05	7.20	8.70
	Follow-up		7.54	-1.38	-2.73	0.09	7.17	8.50
Anxiety	Baseline[b]	[0, 6]	0.00	0.00	0.00	0.00	6.00	6.00
	Session 2		2.19	-2.19	-4.55	2.70	8.29	10.04
	Session 4		4.66	-2.79	-7.15	2.24	8.08	9.77
	Follow-up		7.49	-3.19	-4.71	2.41	8.27	10.19
Depression	Baseline[b]	[0, 6]	0.00	0.00	0.00	0.00	6.00	6.00
	Session 2		4.41	-2.67	-4.41	1.41	7.83	9.24
	Session 4		9.13	-3.36	-5.39	1.17	7.74	9.79
	Follow-up		14.09	-3.82	-5.93	1.28	7.95	10.39
Self-Efficacy	Baseline[b]	[0, 4]	0.00	-0.01	-0.73	0.00	4.00	4.09
	Session 2		0.09	-0.14	-0.66	2.22	5.05	5.91
	Session 4		0.09	-0.18	-1.35	7.01	5.57	6.58
	Follow-up		0.12	-0.28	-1.38	12.18	5.97	7.21
Growth Mindset	Baseline[b]	[0, 4]	0.00	0.00	-0.08	0.00	4.00	4.00
	Session 2		0.21	-0.31	-1.42	1.38	4.92	5.93
	Session 4		0.41	-0.56	-1.29	4.88	5.54	6.76
	Follow-up		0.62	-0.75	-1.83	8.74	5.91	6.78
Optimism	Baseline[b]	[0, 4]	0.01	-0.02	-0.79	0.00	4.00	4.00
	Session 2		1.04	-0.83	-1.94	0.52	4.53	5.05
	Session 4		1.39	-0.96	-1.75	2.40	5.23	6.44
	Follow-up		1.64	-1.11	-2.10	4.76	5.63	7.20
Separate-Level Dataset								
Positive Bias	Baseline[a]	[1, 7]	0.00	1.00	1.00	0.00	6.00	6.00
	Session 1		0.05	0.91	-0.02	0.02	7.06	8.05
	Session 2		0.09	0.78	-0.32	0.34	7.56	9.10
	Session 3		0.09	0.84	-0.80	1.56	8.32	9.59
	Session 4		0.12	0.65	-0.40	5.34	9.18	10.31
	Follow-up		0.18	0.51	-0.96	13.39	10.09	11.58
Negative Bias	Baseline[a]	[1, 7]	0.00	2.00	2.00	0.00	7.00	7.00
	Session 1		0.03	0.94	-0.10	0.01	7.04	8.29

	Session 2		0.26	0.54	-0.36	0.03	6.78	7.88
	Session 3		1.24	-0.14	-0.87	0.03	6.90	7.79
	Session 4		3.19	-0.70	-2.23	0.04	7.13	8.35
	Follow-up		7.72	-1.33	-2.29	0.07	7.11	9.12
Anxiety	Baseline[b]	[0, 6]	0.00	0.00	0.00	0.00	6.00	6.00
	Session 2		2.08	-2.19	-4.54	2.73	8.26	10.83
	Session 4		4.62	-2.69	-4.61	2.30	8.36	11.19
	Follow-up		7.68	-3.19	-5.00	2.53	8.44	10.13
Depression	Baseline[b]	[0, 6]	0.00	0.00	0.00	0.00	6.00	6.00
	Session 2		4.25	-2.69	-5.48	1.44	7.79	9.24
	Session 4		9.18	-3.46	-5.92	1.24	7.87	11.37
	Follow-up		14.39	-3.86	-5.52	1.40	7.91	9.79
Self-Efficacy	Baseline[b]	[0, 4]	0.00	-0.01	-0.40	0.00	4.00	4.06
	Session 2		0.10	-0.14	-0.82	2.11	4.99	5.63
	Session 4		0.09	-0.18	-1.35	7.00	5.60	6.75
	Follow-up		0.13	-0.25	-1.14	12.35	6.02	7.15
Growth Mindset	Baseline[b]	[0, 4]	0.00	0.00	-0.28	0.00	4.00	4.00
	Session 2		0.21	-0.34	-1.53	1.34	4.85	5.53
	Session 4		0.43	-0.63	-1.91	4.80	5.54	6.55
	Follow-up		0.68	-0.84	-2.05	8.70	5.94	7.00
Optimism	Baseline[b]	[0, 4]	0.01	-0.05	-1.02	0.00	4.00	4.00
	Session 2		1.17	-0.86	-2.06	0.50	4.61	5.44
	Session 4		1.47	-1.02	-2.15	2.46	5.29	6.24
	Follow-up		1.70	-1.19	-2.34	4.80	5.68	7.42

[a]Screening.
[b]Pretreatment.

Table S6

Multilevel Modeling Time Effects in Each Condition for the Per-Protocol Sample

Outcome	Phase	Condition	β (*SE*)	*df*	*t*	*p*	*d*
Positive Bias	TX	Positive + Negation	0.37 (0.05)	141.04	7.41	< .001§§	1.75
		Positive	0.46 (0.04)	177.03	12.53	< .001§§	1.89
		50/50 Blocked	0.28 (0.04)	165.04	6.49	< .001§§	1.42
		50/50 Random	0.36 (0.05)	197.03	8.17	< .001§§	2.24
		Neutral Control	0.28 (0.03)	461.01	10.40	< .001§§	1.11
	FU	Positive + Negation	0.49 (0.21)	21.29	2.32	.030	0.40
		Positive	0.36 (0.19)	28.91	1.91	.067	0.39
		50/50 Blocked	0.02 (0.18)	27.80	0.10	.920	0.02
		50/50 Random	0.46 (0.18)	36.92	2.50	.017§	0.37
		Neutral Control	0.16 (0.11)	67.95	1.41	.165	0.14
Negative Bias	TX	Positive + Negation	-0.27 (0.05)	141.04	-5.75	< .001§§	-1.14
		Positive	-0.28 (0.03)	177.03	-8.74	< .001§§	-1.32
		50/50 Blocked	-0.19 (0.04)	165.04	-4.33	< .001§§	-0.76
		50/50 Random	-0.17 (0.04)	197.03	-4.25	< .001§§	-0.76
		Neutral Control	-0.21 (0.03)	461.01	-8.42	< .001§§	-0.83
	FU	Positive + Negation	-0.30 (0.20)	20.48	-1.52	.144	-0.28
		Positive	-0.29 (0.17)	29.75	-1.72	.096	-0.29
		50/50 Blocked	-0.08 (0.16)	29.77	-0.46	.646	-0.07
		50/50 Random	-0.43 (0.16)	34.56	-2.62	.013§	-0.35
		Neutral Control	-0.15 (0.10)	71.97	-1.48	.143	-0.14
Anxiety	TX	Positive + Negation	-0.34 (0.07)	69.08	-4.78	< .001§§	-0.82
		Positive	-0.14 (0.06)	87.07	-2.18	.032	-0.29
		50/50 Blocked	-0.11 (0.07)	81.07	-1.55	.125	-0.26
		50/50 Random	-0.14 (0.07)	97.06	-1.95	.054	-0.29
		Neutral Control	-0.22 (0.04)	229.03	-5.49	< .001§§	-0.46
	FU	Positive + Negation	-0.19 (0.31)	18.98	-0.61	.548	-0.11
		Positive	-0.12 (0.33)	27.37	-0.36	.719	-0.06
		50/50 Blocked	0.12 (0.35)	32.76	0.36	.724	0.07
		50/50 Random	-0.60 (0.32)	35.08	-1.86	.071	-0.35
		Neutral Control	-0.09 (0.19)	81.25	-0.46	.650	-0.05
Depression	TX	Positive + Negation	-0.24 (0.08)	69.08	-2.92	.005§	-0.43
		Positive	-0.14 (0.06)	87.07	-2.63	.010§	-0.29
		50/50 Blocked	-0.17 (0.07)	81.07	-2.38	.020§	-0.35
		50/50 Random	-0.22 (0.07)	97.06	-3.27	.001§§	-0.45
		Neutral Control	-0.16 (0.04)	229.03	-4.17	< .001§§	-0.33
	FU	Positive + Negation	-0.11 (0.30)	21.70	-0.35	.727	-0.05
		Positive	-0.08 (0.31)	29.80	-0.27	.793	-0.04
		50/50 Blocked	-0.09 (0.32)	31.65	-0.28	.782	-0.04
		50/50 Random	-0.19 (0.32)	37.72	-0.61	.545	-0.11
		Neutral Control	-0.04 (0.18)	74.78	-0.23	.816	-0.02
Self-Efficacy	TX	Positive + Negation[a]	0.13 (0.02)	69.08	5.29	< .001§§	0.51
		Positive[a]	0.14 (0.02)	87.07	5.89	< .001§§	0.68
		50/50 Blocked	0.17 (0.03)	81.07	5.35	< .001§§	0.72

		50/50 Random	0.16 (0.03)	97.06	4.74	< .001§§	0.70
		Neutral Control	0.11 (0.02)	229.03	6.14	< .001§§	0.44
	FU	Positive + Negation	0.09 (0.13)	17.97	0.71	.488	0.10
		Positive	0.15 (0.12)	29.82	1.34	.189	0.26
		50/50 Blocked	0.13 (0.12)	29.60	1.12	.272	0.16
		50/50 Random	0.04 (0.13)	37.28	0.31	.757	0.05
		Neutral Control	0.10 (0.07)	67.95	1.46	.150	0.12
Growth	TX	Positive + Negation	0.11 (0.04)	69.08	2.55	.013§	0.55
Mindset		Positive	0.09 (0.03)	87.07	3.42	.001§§	0.38
		50/50 Blocked[a]	0.07 (0.03)	81.07	2.65	.010§	0.41
		50/50 Random	0.12 (0.03)	97.06	3.88	< .001§§	0.47
		Neutral Control	0.06 (0.02)	229.03	3.09	.002§§	0.25
	FU	Positive + Negation	0.27 (0.15)	19.97	1.84	.081	0.29
		Positive	0.18 (0.12)	30.27	1.46	.153	0.18
		50/50 Blocked	0.19 (0.13)	25.17	1.51	.143	0.25
		50/50 Random	0.06 (0.13)	36.81	0.44	.661	0.06
		Neutral Control	0.06 (0.07)	67.19	0.86	.396	0.06
Optimism	TX	Positive + Negation	0.17 (0.03)	69.08	5.74	< .001§§	0.68
		Positive	0.13 (0.03)	87.07	4.52	< .001§§	0.52
		50/50 Blocked	0.11 (0.03)	81.07	3.67	< .001§§	0.44
		50/50 Random	0.08 (0.04)	97.06	2.14	.035	0.32
		Neutral Control	0.10 (0.02)	229.03	5.40	< .001§§	0.37
	FU	Positive + Negation	0.18 (0.14)	18.75	1.31	.207	0.20
		Positive	0.25 (0.12)	27.30	2.05	.051	0.27
		50/50 Blocked	0.20 (0.14)	29.83	1.45	.157	0.21
		50/50 Random	0.30 (0.12)	36.29	2.56	.015§	0.28
		Neutral Control	0.04 (0.07)	66.49	0.55	.586	0.04

Note. Separate models were fit for each outcome, phase, and condition. Every model included the fixed effect of time (shown here). Treatment phase models included a random intercept and random slope for time (except where noted); follow-up phase models included only a random intercept. The Separate-Level Dataset, with condition coded in five levels (Positive Prospection + Negation, Positive Prospection, 50/50 Blocked, 50/50 Random, Neutral Control), was used. To correct for multiple comparisons among models in Table S3, the Bonferroni-corrected alpha level is .025. TX = treatment; FU = follow-up.

[a]Because models including a random intercept and random slope for time did not converge, we removed the random slope for time, leaving only the fixed effect of time and a random intercept.

§$p < .025$

§§$p < .005$

Table S7

Multilevel Modeling Fixed Condition × Time Interaction and Simple Time Effects for the Per-Protocol Sample

Outcome	Phase	Effect	β (*SE*)	*df*	*t*	*p*	*d*, 97.5% CI
Positive Bias	TX	(Both Positive vs. Neutral Control) × Time	0.14 (0.04)	1151.01	3.36	.001§§	0.58 [0.19, 0.97]
		$Time_{Both\ Positive}$	0.42 (0.03)	321.02	13.89	< .001§§	1.84
		$Time_{Neutral\ Control}$	0.28 (0.03)	461.01	10.40	< .001§§	1.11
		(Both Positive vs. Both 50/50) × Time	0.09 (0.04)	1151.01	2.04	.042	0.44 [-0.04, 0.92]
		(Both 50/50 vs. Neutral Control) × Time	0.05 (0.04)	1151.01	1.26	.209	0.23 [-0.18, 0.63]
		(Positive + Negation vs. Positive) × Time	-0.09 (0.06)	1149.01	-1.41	.159	-0.40 [-1.03, 0.23]
		(50/50 Blocked vs. 50/50 Random) × Time	-0.08 (0.06)	1149.01	-1.36	.175	-0.45 [-1.20, 0.30]
	FU	(Both Positive vs. Neutral Control) × Time	0.25 (0.18)	165.95	1.40	.162	0.23 [-0.14, 0.59]
		(Both Positive vs. Both 50/50) × Time	0.15 (0.18)	202.09	0.86	.390	0.13 [-0.22, 0.49]
		(Both 50/50 vs. Neutral Control) × Time	0.10 (0.17)	179.65	0.59	.557	0.08 [-0.24, 0.41]
		(Positive + Negation vs. Positive) × Time	0.13 (0.28)	165.88	0.48	.633	0.12 [-0.46, 0.71]
		(50/50 Blocked vs. 50/50 Random) × Time	-0.44 (0.25)	181.02	-1.75	.082	-0.37 [-0.84, 0.11]
Negative Bias	TX	(Both Positive vs. Neutral Control) × Time	-0.07 (0.04)	1151.01	-1.77	.078	-0.28 [-0.64, 0.08]
		(Both Positive vs. Both 50/50) × Time	-0.10 (0.04)	1151.01	-2.44	.015§	-0.43 [-0.83, -0.03]
		$Time_{Both\ Positive}$	-0.28 (0.03)	321.02	-10.09	< .001§§	-1.22
		$Time_{Both\ 50/50}$	-0.18 (0.03)	365.02	-6.09	< .001§§	-0.76
		(Both 50/50 vs. Neutral Control) × Time	0.03 (0.04)	1151.01	0.83	.405	0.13 [-0.22, 0.47]
		(Positive + Negation vs. Positive) × Time	0.00 (0.06)	1149.01	0.01	.991	0.00 [-0.59, 0.60]
		(50/50 Blocked vs. 50/50 Random) × Time	-0.02 (0.06)	1149.01	-0.35	.729	-0.08 [-0.62, 0.45]
	FU	(Both Positive vs. Neutral Control) × Time	-0.13 (0.17)	160.97	-0.77	.442	-0.12 [-0.46, 0.23]
		(Both Positive vs. Both 50/50) × Time	-0.02 (0.17)	182.79	-0.11	.911	-0.02 [-0.36, 0.32]
		(Both 50/50 vs. Neutral Control) × Time	-0.11 (0.15)	193.65	-0.72	.471	-0.10 [-0.39, 0.20]
		(Positive + Negation vs. Positive) × Time	-0.02 (0.27)	140.72	-0.07	.946	-0.02 [-0.6, 0.57]
		(50/50 Blocked vs. 50/50 Random) × Time	0.35 (0.23)	185.86	1.54	.125	0.30 [-0.14, 0.74]
Anxiety	TX	(Both Positive vs. Neutral Control) × Time	-0.01 (0.07)	573.01	-0.17	.869	-0.02 [-0.33, 0.29]
		(Both Positive vs. Both 50/50) × Time	-0.10 (0.07)	573.01	-1.47	.142	-0.22 [-0.55, 0.11]
		(Both 50/50 vs. Neutral Control) × Time	0.09 (0.06)	573.01	1.43	.152	0.19 [-0.11, 0.49]
		(Positive + Negation vs. Positive) × Time	-0.20 (0.10)	571.01	-2.01	.045	-0.44 [-0.93, 0.05]
		(50/50 Blocked vs. 50/50 Random) × Time	0.03 (0.09)	571.01	0.29	.774	0.06 [-0.40, 0.51]

	FU	(Both Positive vs. Neutral Control) × Time	-0.06 (0.31)	177.85	-0.18	.858	-0.03 [-0.43, 0.36]
		(Both Positive vs. Both 50/50) × Time	0.12 (0.33)	176.60	0.38	.708	0.07 [-0.35, 0.49]
		(Both 50/50 vs. Neutral Control) × Time	-0.18 (0.28)	209.38	-0.63	.532	-0.10 [-0.47, 0.26]
		(Positive + Negation vs. Positive) × Time	-0.07 (0.50)	150.00	-0.14	.891	-0.04 [-0.67, 0.59]
		(50/50 Blocked vs. 50/50 Random) × Time	0.72 (0.43)	200.20	1.69	.093	0.42 [-0.14, 0.98]
Depression	TX	(Both Positive vs. Neutral Control) × Time	-0.03 (0.06)	573.01	-0.42	.678	-0.05 [-0.33, 0.23]
		(Both Positive vs. Both 50/50) × Time	0.01 (0.07)	573.01	0.16	.872	0.02 [-0.27, 0.31]
		(Both 50/50 vs. Neutral Control) × Time	-0.04 (0.06)	573.01	-0.61	.545	-0.07 [-0.35, 0.20]
		(Positive + Negation vs. Positive) × Time	-0.09 (0.10)	571.01	-0.96	.339	-0.18 [-0.59, 0.24]
		(50/50 Blocked vs. 50/50 Random) × Time	0.05 (0.09)	571.01	0.60	.552	0.11 [-0.31, 0.53]
	FU	(Both Positive vs. Neutral Control) × Time	-0.02 (0.30)	171.88	-0.06	.953	-0.01 [-0.37, 0.35]
		(Both Positive vs. Both 50/50) × Time	0.09 (0.31)	175.09	0.29	.772	0.05 [-0.32, 0.41]
		(Both 50/50 vs. Neutral Control) × Time	-0.11 (0.27)	202.37	-0.40	.693	-0.06 [-0.38, 0.27]
		(Positive + Negation vs. Positive) × Time	-0.03 (0.45)	177.54	-0.06	.955	-0.01 [-0.55, 0.52]
		(50/50 Blocked vs. 50/50 Random) × Time	0.11 (0.41)	190.40	0.26	.799	0.05 [-0.43, 0.54]
Self-Efficacy	TX	(Both Positive vs. Neutral Control) × Time	0.03 (0.03)	573.01	1.09	.277	0.12 [-0.13, 0.38]
		(Both Positive vs. Both 50/50) × Time	-0.03 (0.03)	573.01	-0.87	.382	-0.11 [-0.39, 0.17]
		(Both 50/50 vs. Neutral Control) × Time	0.05 (0.03)	573.01	2.08	.038	0.23 [-0.02, 0.48]
		(Positive + Negation vs. Positive) × Time	-0.02 (0.04)	571.01	-0.42	.672	-0.08 [-0.49, 0.33]
		(50/50 Blocked vs. 50/50 Random) × Time	0.01 (0.04)	571.01	0.25	.804	0.04 [-0.34, 0.43]
	FU	(Both Positive vs. Neutral Control) × Time	0.02 (0.12)	134.82	0.20	.844	0.03 [-0.31, 0.37]
		(Both Positive vs. Both 50/50) × Time	0.05 (0.13)	142.82	0.38	.708	0.06 [-0.30, 0.41]
		(Both 50/50 vs. Neutral Control) × Time	-0.02 (0.11)	172.17	-0.21	.832	-0.03 [-0.32, 0.26]
		(Positive + Negation vs. Positive) × Time	-0.06 (0.18)	149.91	-0.34	.736	-0.08 [-0.62, 0.46]
		(50/50 Blocked vs. 50/50 Random) × Time	0.09 (0.16)	175.27	0.57	.568	0.11 [-0.33, 0.55]
Growth Mindset	TX	(Both Positive vs. Neutral Control) × Time	0.04 (0.03)	573.01	1.46	.144	0.19 [-0.10, 0.48]
		(Both Positive vs. Both 50/50) × Time	0.00 (0.03)	573.01	0.00	.998	0.00 [-0.31, 0.31]
		(Both 50/50 vs. Neutral Control) × Time	0.04 (0.03)	573.01	1.52	.129	0.19 [-0.09, 0.47]
		(Positive + Negation vs. Positive) × Time	0.02 (0.05)	571.01	0.45	.657	0.09 [-0.37, 0.55]
		(50/50 Blocked vs. 50/50 Random) × Time	-0.05 (0.04)	571.01	-1.16	.247	-0.22 [-0.64, 0.21]
	FU	(Both Positive vs. Neutral Control) × Time	0.16 (0.12)	167.15	1.38	.171	0.17 [-0.11, 0.45]
		(Both Positive vs. Both 50/50) × Time	0.11 (0.12)	174.25	0.87	.386	0.12 [-0.18, 0.41]
		(Both 50/50 vs. Neutral Control) × Time	0.06 (0.11)	184.47	0.51	.614	0.06 [-0.21, 0.33]

		(Positive + Negation vs. Positive) × Time	0.09 (0.19)	143.94	0.46	.649	0.09 [-0.36, 0.54]
		(50/50 Blocked vs. 50/50 Random) × Time	0.14 (0.17)	167.69	0.81	.422	0.16 [-0.28, 0.60]
Optimism	TX	(Both Positive vs. Neutral Control) × Time	0.05 (0.03)	573.01	1.74	.082	0.20 [-0.06, 0.46]
		(Both Positive vs. Both 50/50) × Time	0.05 (0.03)	573.01	1.69	.092	0.21 [-0.07, 0.48]
		(Both 50/50 vs. Neutral Control) × Time	0.00 (0.03)	573.01	-0.04	.969	0.00 [-0.25, 0.24]
		(Positive + Negation vs. Positive) × Time	0.04 (0.05)	571.01	0.87	.384	0.16 [-0.25, 0.57]
		(50/50 Blocked vs. 50/50 Random) × Time	0.03 (0.04)	571.01	0.72	.471	0.12 [-0.26, 0.50]
	FU	(Both Positive vs. Neutral Control) × Time	0.16 (0.12)	143.70	1.28	.201	0.17 [-0.13, 0.46]
		(Both Positive vs. Both 50/50) × Time	-0.05 (0.13)	149.03	-0.42	.673	-0.06 [-0.36, 0.25]
		(Both 50/50 vs. Neutral Control) × Time	0.21 (0.12)	156.49	1.84	.068	0.22 [-0.05, 0.48]
		(Positive + Negation vs. Positive) × Time	-0.07 (0.19)	137.66	-0.35	.728	-0.07 [-0.55, 0.41]
		(50/50 Blocked vs. 50/50 Random) × Time	-0.10 (0.17)	178.75	-0.63	.531	-0.10 [-0.48, 0.27]

Note. Separate models were fit for each outcome, phase, and reference group. Each model contained the fixed effects of condition, time, and the Condition × Time interaction. Treatment phase models included a random intercept and random slope for time; follow-up phase models included only a random intercept. The latter level of the dummy-coded condition factor in each interaction effect is the reference group. Simple time effects are shown only for significant interactions. Significance is based on a Bonferroni-corrected alpha level of .025 (.05/2 given two orthogonal interactions per dataset). The Combined-Level Dataset, with condition coded in three levels (Both Positive, Both 50/50, Neutral Control), was used to test interactions contrasting Both Positive with Neutral Control, Both Positive with Both 50/50, and Both 50/50 with Neutral Control. The Separate-Level Dataset, with condition coded in five levels (Positive Prospection + Negation, Positive Prospection, 50/50 Blocked, 50/50 Random, Neutral Control), was used to test interaction effects contrasting Positive Prospection + Negation with Positive Prospection and 50/50 Blocked with 50/50 Random. TX = treatment; FU = follow-up.

$^{\S}p < .025$

$^{\S\S}p < .005$

Table S8

Exploratory Multilevel Modeling Time Effects in Each Condition on Anxiety and Depression Composite, by Sample

Phase	Condition	β (*SE*)	*df*	*t*	*d*
		Intent-To-Treat Sample			
TX	Positive + Negation	-0.54 (0.10)	77.73	-5.26	-0.67
	Positive	-0.35 (0.09)	89.26	-3.87	-0.41
	50/50 Blocked	-0.33 (0.09)	99.86	-3.71	-0.44
	50/50 Random	-0.41 (0.09)	103.50	-4.51	-0.49
	Neutral Control	-0.41 (0.06)	182.08	-7.23	-0.51
FU	Positive + Negation	-0.46 (0.31)	61.76	-1.47	-0.14
	Positive	-0.30 (0.29)	78.67	-1.03	-0.09
	50/50 Blocked	-0.21 (0.33)	71.03	-0.65	-0.06
	50/50 Random	-0.52 (0.31)	80.84	-1.70	-0.16
	Neutral Control	-0.29 (0.21)	133.26	-1.41	-0.09
		Per-Protocol Sample			
TX	Positive + Negation	-0.58 (0.13)	69.08	-4.33	-0.66
	Positive	-0.28 (0.10)	87.07	-2.74	-0.33
	50/50 Blocked	-0.28 (0.12)	81.07	-2.32	-0.36
	50/50 Random	-0.36 (0.12)	97.06	-3.01	-0.45
	Neutral Control	-0.38 (0.06)	229.03	-6.13	-0.47
FU	Positive + Negation	-0.30 (0.52)	19.83	-0.56	-0.08
	Positive	-0.20 (0.56)	28.67	-0.36	-0.06
	50/50 Blocked	0.04 (0.60)	33.64	0.06	0.01
	50/50 Random	-0.79 (0.58)	37.13	-1.38	-0.25
	Neutral Control	-0.13 (0.31)	77.03	-0.41	-0.04

Table S9

Exploratory Multilevel Modeling Fixed Condition × Time Interaction Effects on Anxiety and Depression Composite, by Sample

Phase	Effect	β (*SE*)	*df*	*t*	*d*, 97.5% CI
	Intent-To-Treat Sample				
TX	(Both Positive vs. Neutral Control) × Time	-0.02 (0.09)	221.96	-0.22	-0.02 [-0.27, 0.23]
	(Both Positive vs. Both 50/50) × Time	-0.06 (0.09)	205.07	-0.68	-0.08 [-0.34, 0.18]
	(Both 50/50 vs. Neutral Control) × Time	0.04 (0.09)	239.91	0.49	0.05 [-0.19, 0.30]
	(Positive + Negation vs. Positive) × Time	-0.20 (0.13)	196.03	-1.46	-0.24 [-0.60, 0.13]
	(50/50 Blocked vs. 50/50 Random) × Time	0.08 (0.13)	226.70	0.62	0.10 [-0.26, 0.46]
FU	(Both Positive vs. Neutral Control) × Time	-0.06 (0.31)	233.29	-0.20	-0.02 [-0.24, 0.20]
	(Both Positive vs. Both 50/50) × Time	0.02 (0.31)	227.06	0.05	0.00 [-0.21, 0.22]
	(Both 50/50 vs. Neutral Control) × Time	-0.08 (0.30)	267.09	-0.26	-0.02 [-0.24, 0.19]
	(Positive + Negation vs. Positive) × Time	-0.16 (0.44)	235.19	-0.36	-0.05 [-0.35, 0.25]
	(50/50 Blocked vs. 50/50 Random) × Time	0.31 (0.43)	251.60	0.71	0.09 [-0.21, 0.40]
	Per-Protocol Sample				
TX	(Both Positive vs. Neutral Control) × Time	-0.04 (0.11)	573.01	-0.34	-0.04 [-0.33, 0.24]
	(Both Positive vs. Both 50/50) × Time	-0.09 (0.11)	573.01	-0.81	-0.11 [-0.41, 0.19]
	(Both 50/50 vs. Neutral Control) × Time	0.05 (0.10)	573.01	0.53	0.07 [-0.22, 0.35]
	(Positive + Negation vs. Positive) × Time	-0.29 (0.16)	571.01	-1.79	-0.34 [-0.76, 0.09]
	(50/50 Blocked vs. 50/50 Random) × Time	0.08 (0.15)	571.01	0.52	0.10 [-0.33, 0.53]
FU	(Both Positive vs. Neutral Control) × Time	-0.07 (0.53)	181.02	-0.14	-0.02 [-0.38, 0.34]
	(Both Positive vs. Both 50/50) × Time	0.21 (0.55)	178.77	0.38	0.06 [-0.31, 0.43]
	(Both 50/50 vs. Neutral Control) × Time	-0.29 (0.49)	205.44	-0.59	-0.09 [-0.43, 0.25]
	(Positive + Negation vs. Positive) × Time	-0.10 (0.84)	162.22	-0.11	-0.03 [-0.57, 0.52]
	(50/50 Blocked vs. 50/50 Random) × Time	0.83 (0.73)	203.12	1.14	0.25 [-0.25, 0.75]

Table S10

Index of Manuscript Content by R Script and Code Section

Manuscript Section / Content	R Script / Code Section
Section Participants and Design	
Data collection start date	2a_cleaning_data.R Analyze participant flow (Part 3)
Size of per-protocol sample	2a_cleaning_data.R Define analysis samples
Section Outcome Measures	
McDonald's omega total	5_baseline_internal_consistency.R Compute McDonald's omega total
Section Missing Data Handling	
Item-level missingness	6_item-level_missingness.R
Scale-level missingness	7a_table s1_scale-level_missingness_auxiliary_variables_MI.R Compute proportions of scale-level missing data
Tables 1-2	8_longitudinal_analyses.R Longitudinal analyses with 5 conditions Longitudinal analyses with 3 conditions
Figure 1	2a_cleaning_data.R Analyze participant flow (Part 1) Analyze participant flow (Part 2) Analyze participant flow (Part 3)
Section S3.1 Open Data	1_get_raw_data.rtf
Section S1.7 Item Selection and Modification	0_item_selection.R
Section S1.10 Baseline Demographics Differences	3_baseline_demographics_differences.R
Section S1.11 Auxiliary Variables	7a_table s1_scale-level_missingness_auxiliary_variables_MI.R Test for auxiliary demographic variables
Table S1	7a_table s1_scale-level_missingness_auxiliary_variables_MI.R Generate Table S1
Table S2	2c_further_cleaning_demographics.R Further clean demographic data and compute descriptives
Table S3	5_baseline_internal_consistency.R

	Compute Cronbach's alpha
Table S4	5_baseline_internal_consistency.R
	Positive expectancy bias (incl. factor analysis)
	Negative expectancy bias (incl. factor analysis)
Table S5	8_longitudinal_analyses.R
	Investigate imputed data distributions and implausible values
Tables S6-S7	8_longitudinal_analyses.R
	Longitudinal analyses with 5 conditions
	Longitudinal analyses with 3 conditions
Tables S8-S9	9_longitudinal_analyses_exploratory.R
	Longitudinal analyses with 5 conditions
	Longitudinal analyses with 3 conditions
Figures S1-S7	7a_table s1_scale-level_missingness_auxiliary_variables_MI.R
	Investigate distributions of data at each time point before imputation
	8_longitudinal_analyses.R
	Investigate imputed data distributions and implausible values

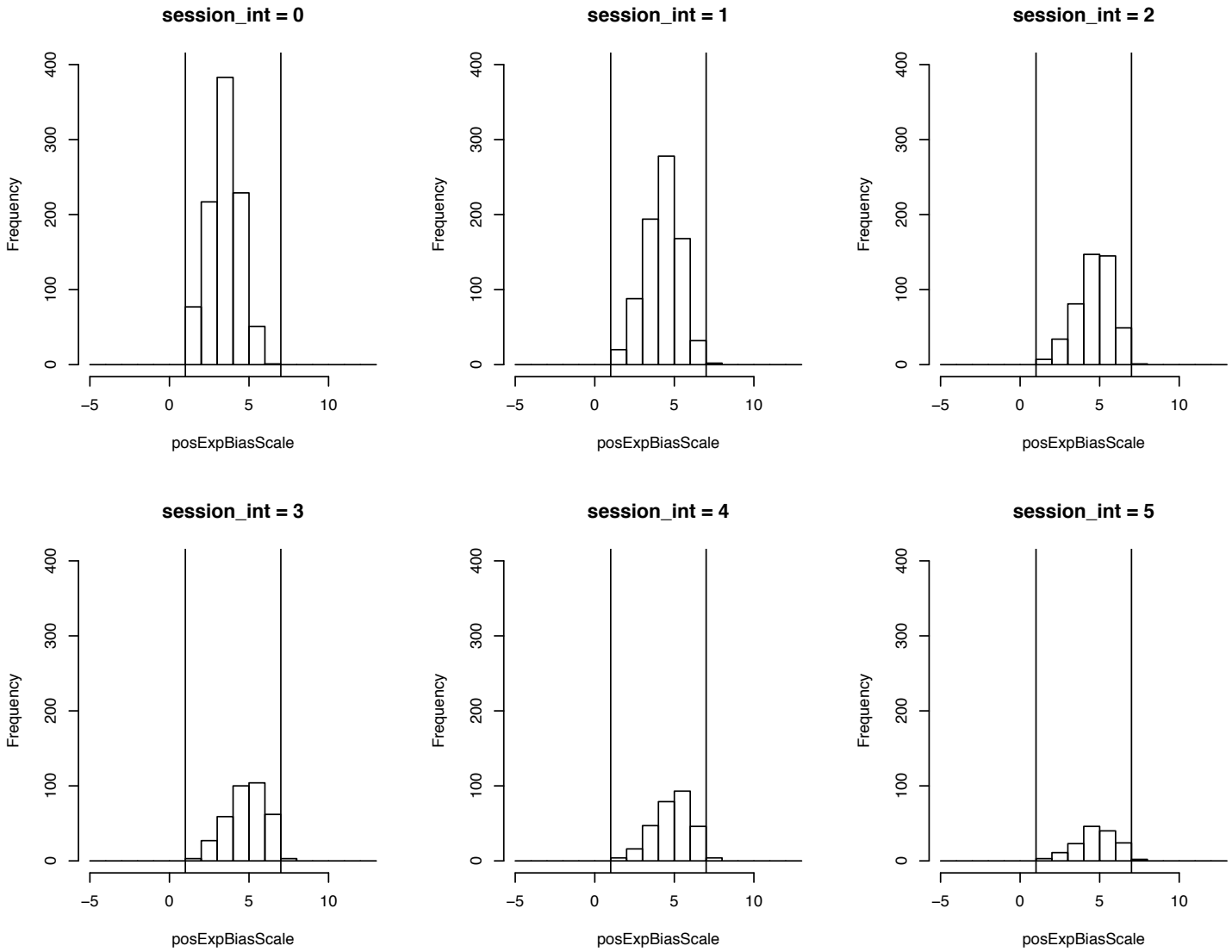

Figure S1.1. Distribution of positive expectancy bias scores at each time point before multiple imputation. Vertical lines show range of plausible values (cells are left-closed).

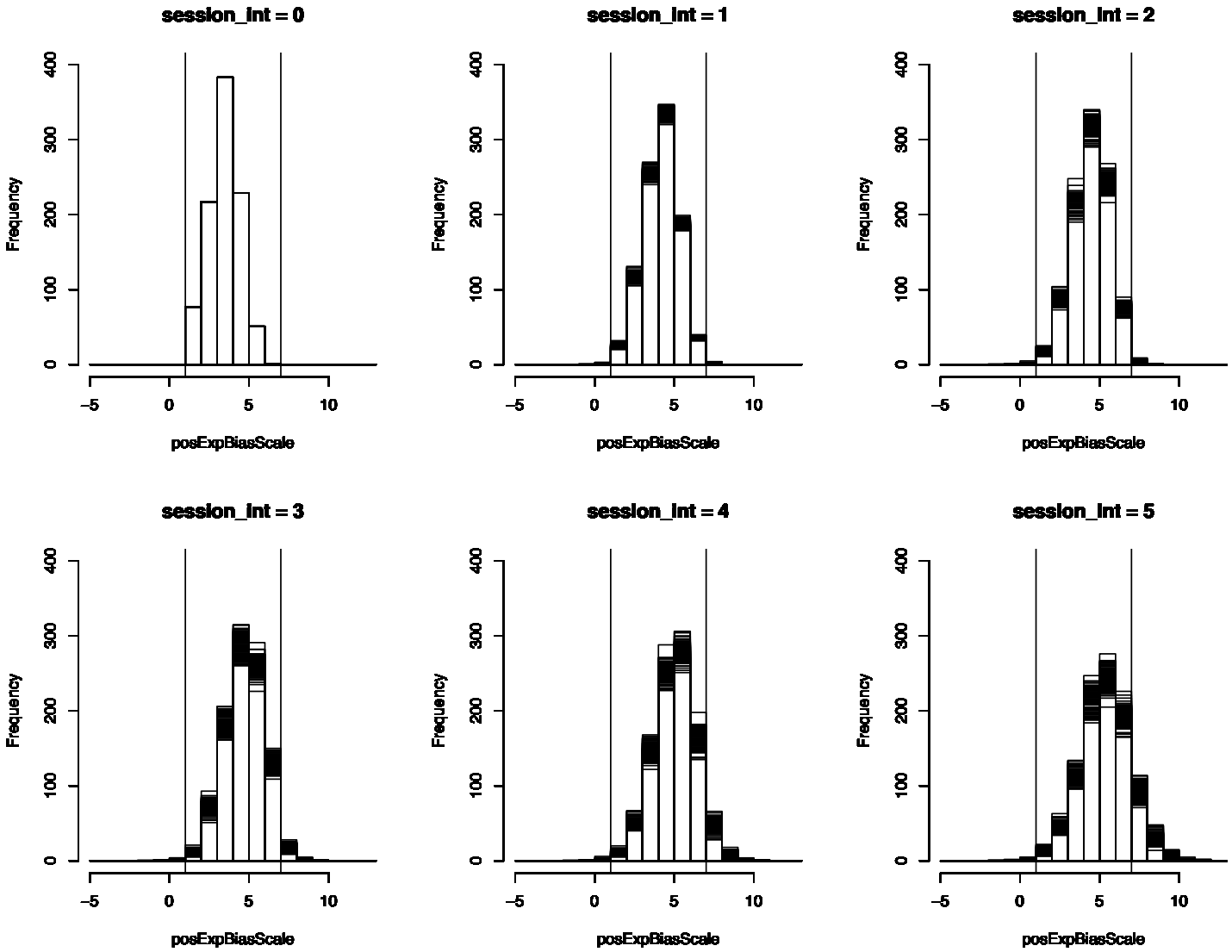

Figure S1.2. Overlaid distributions of positive expectancy bias scores at each time point for 100 multiple imputations of Combined-Level Dataset. Vertical lines show range of plausible values (cells are left-closed).

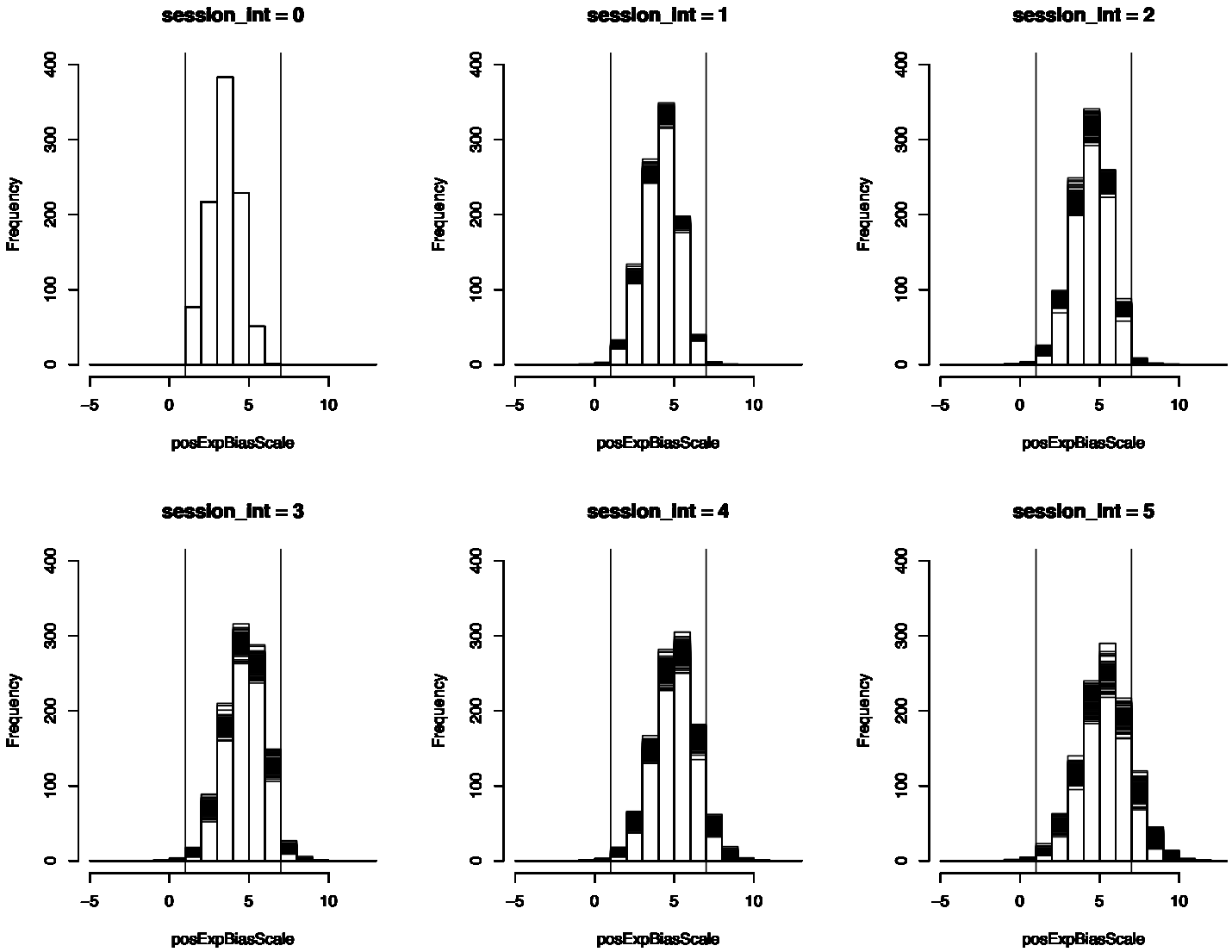

Figure S1.3. Overlaid distributions of positive expectancy bias scores at each time point for 100 multiple imputations of Separate-Level Dataset. Vertical lines show range of plausible values (cells are left-closed).

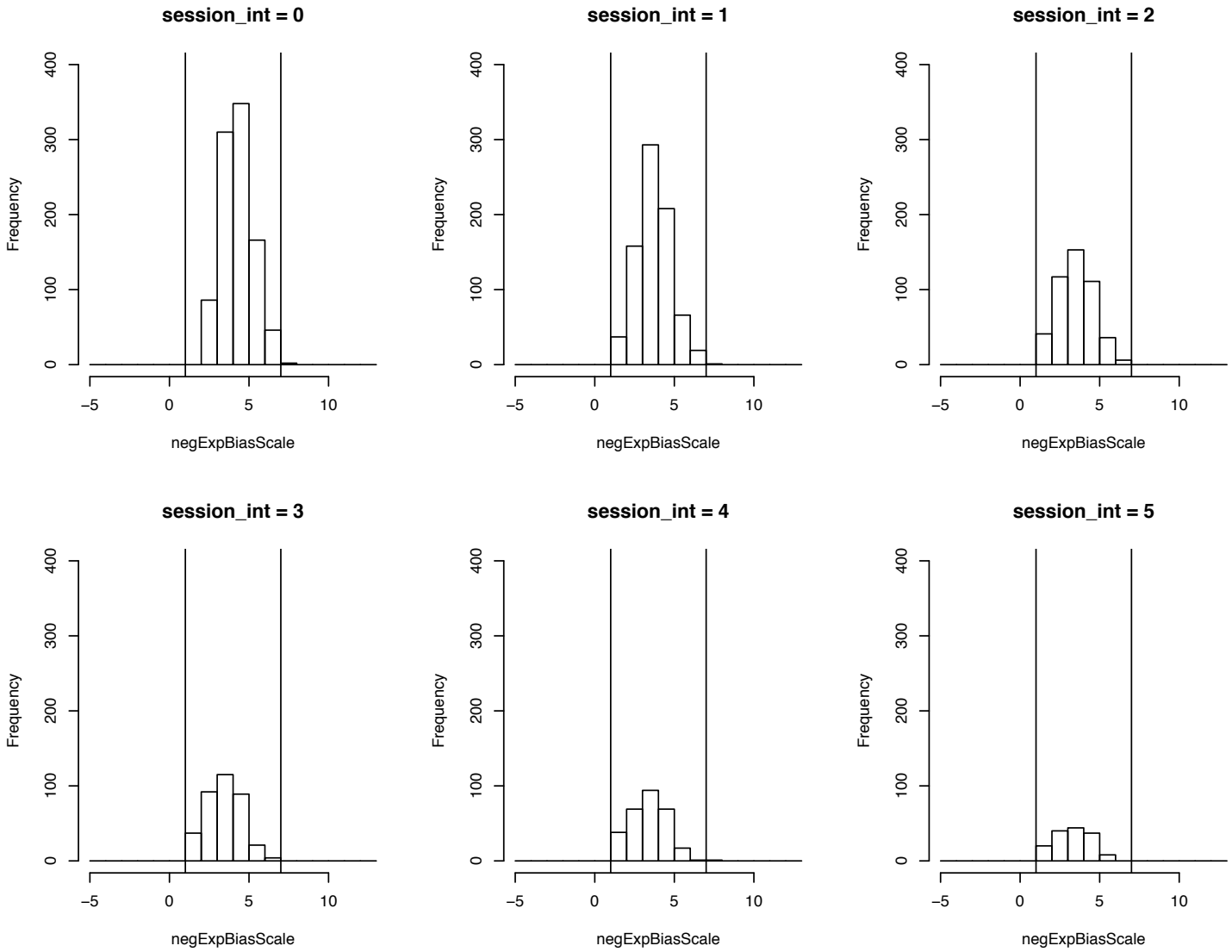

Figure S2.1. Distribution of negative expectancy bias scores at each time point before multiple imputation. Vertical lines show range of plausible values (cells are left-closed).

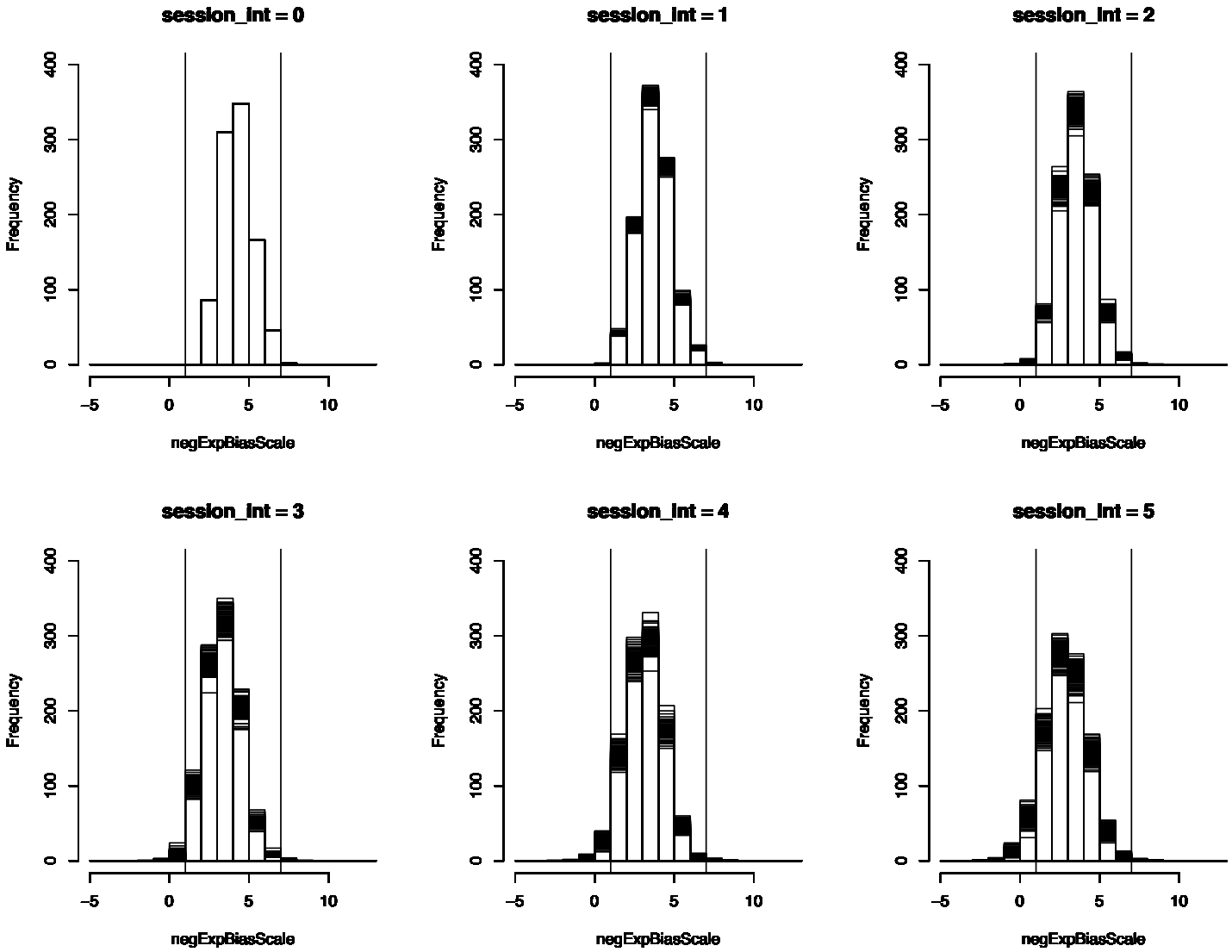

Figure S2.2. Overlaid distributions of negative expectancy bias scores at each time point for 100 multiple imputations of Combined-Level Dataset. Vertical lines show range of plausible values (cells are left-closed).

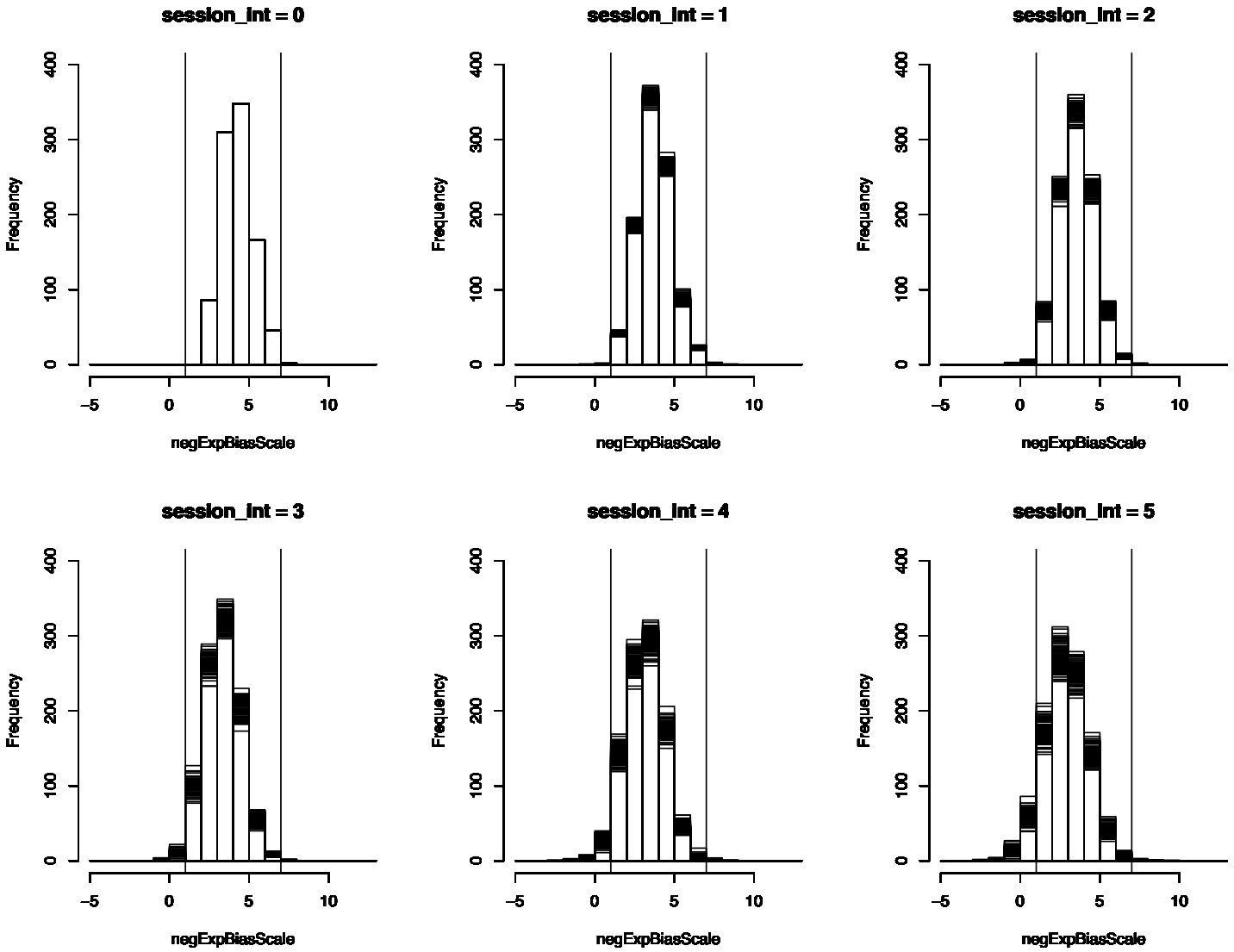

Figure S2.3. Overlaid distributions of negative expectancy bias scores at each time point for 100 multiple imputations of Separate-Level Dataset. Vertical lines show range of plausible values (cells are left-closed).

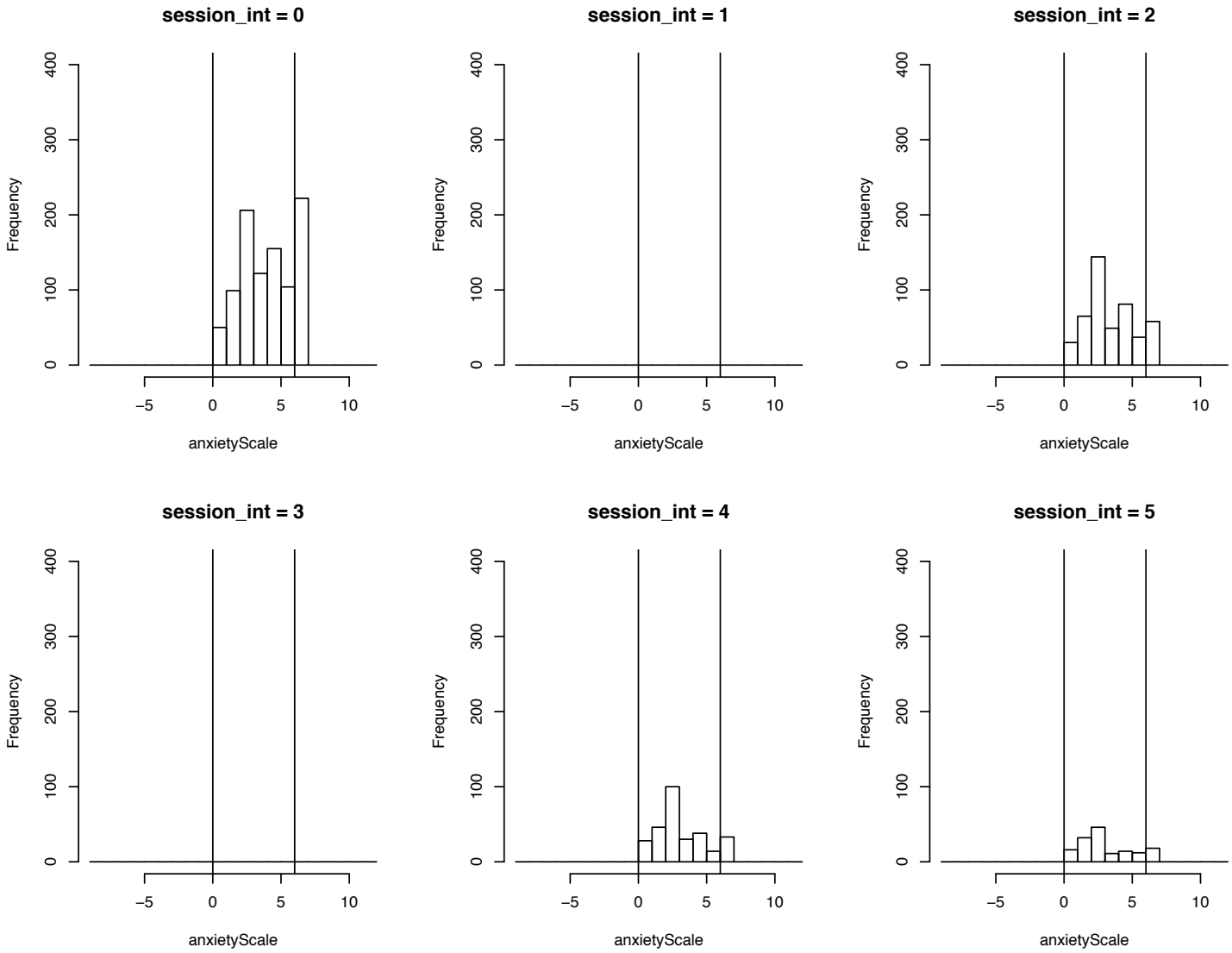

Figure S3.1. Distribution of anxiety scores at each time point before multiple imputation. Vertical lines show range of plausible values (cells are left-closed). Anxiety was not assessed at Sessions 1 or 3.

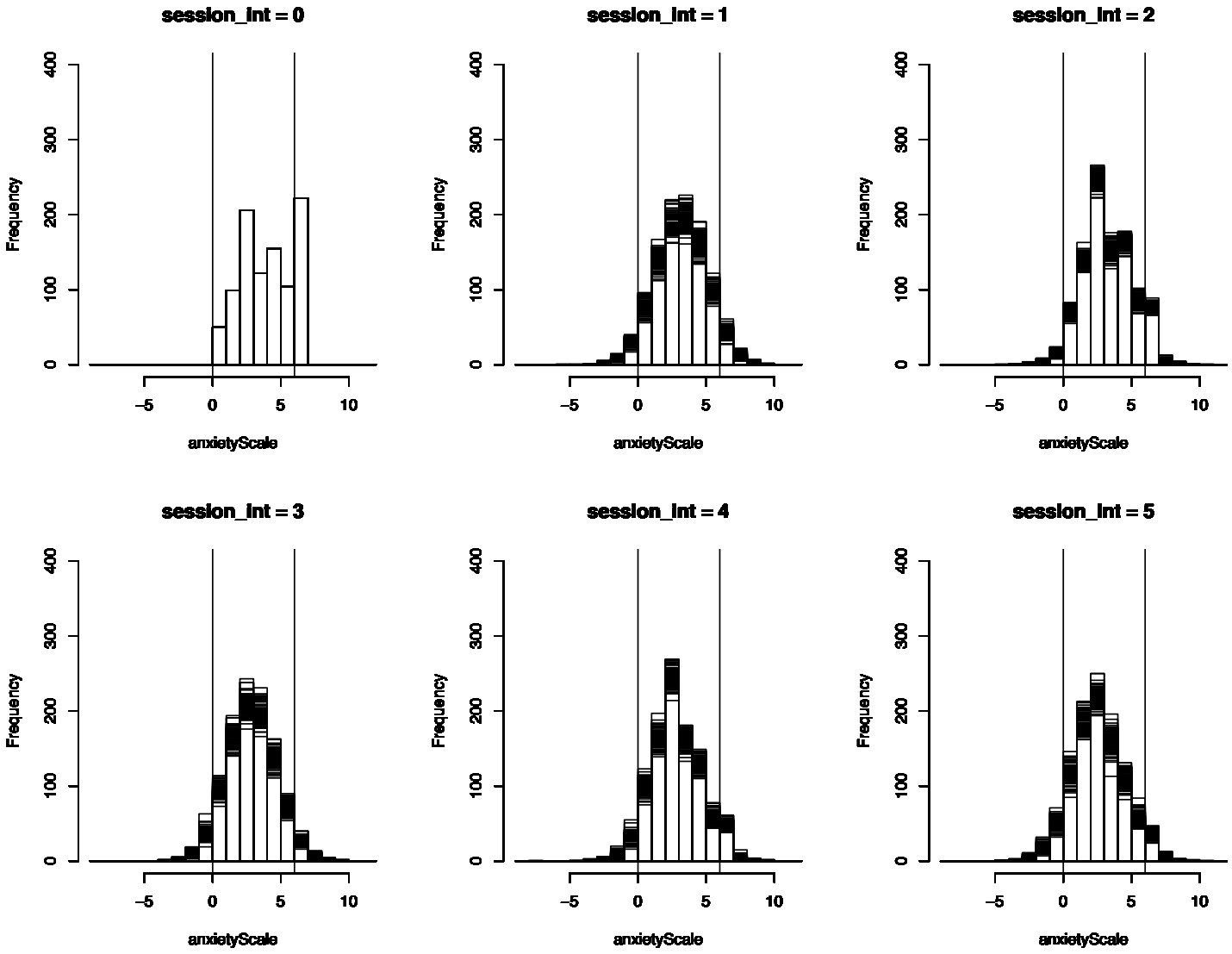

Figure S3.2. Overlaid distributions of anxiety scores at each time point for 100 multiple imputations of Combined-Level Dataset. Vertical lines show range of plausible values (cells are left-closed). Imputed scores for Sessions 1 and 3 were not analyzed.

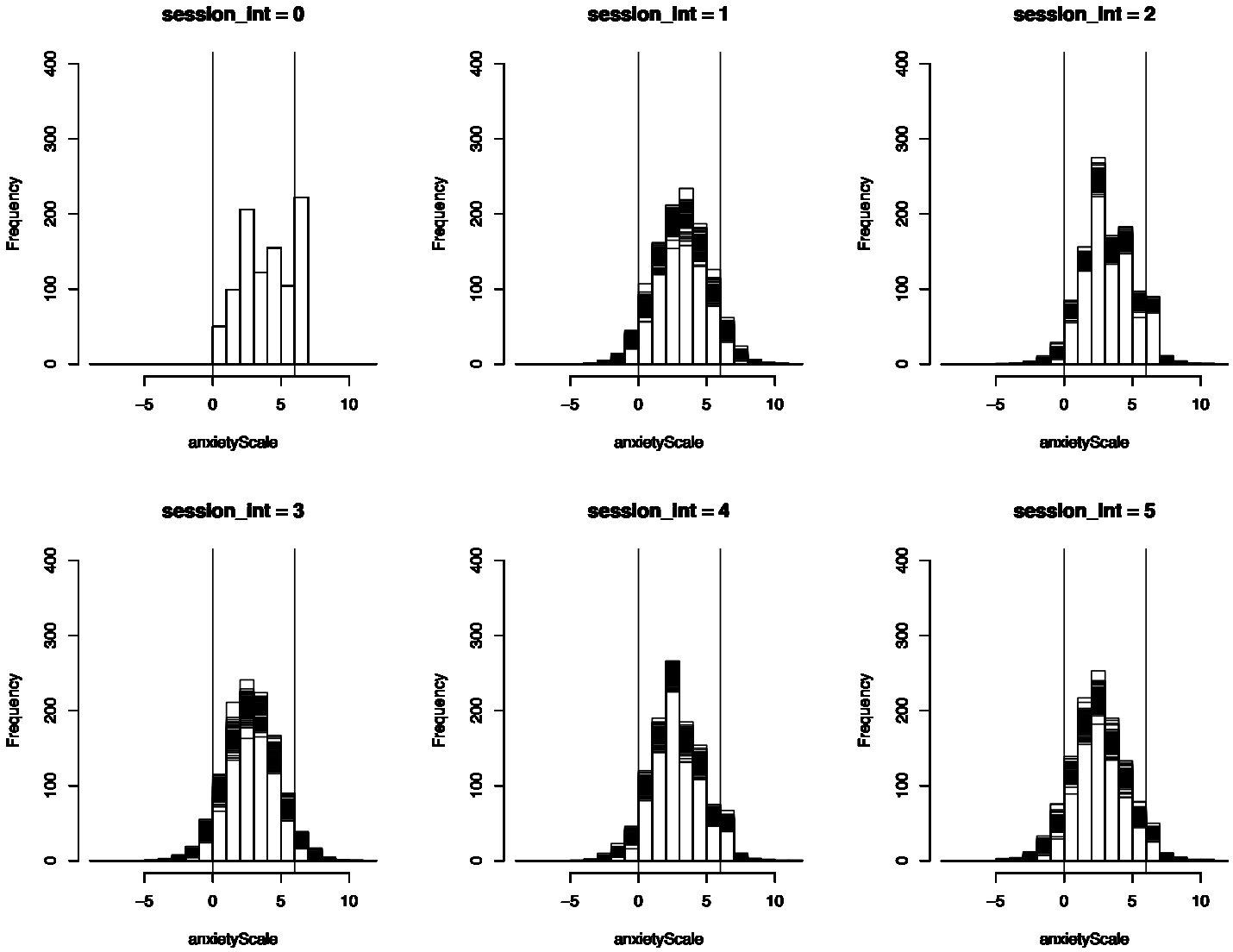

Figure S3.3. Overlaid distributions of anxiety scores at each time point for 100 multiple imputations of Separate-Level Dataset. Vertical lines show range of plausible values (cells are left-closed). Imputed scores for Sessions 1 and 3 were not analyzed.

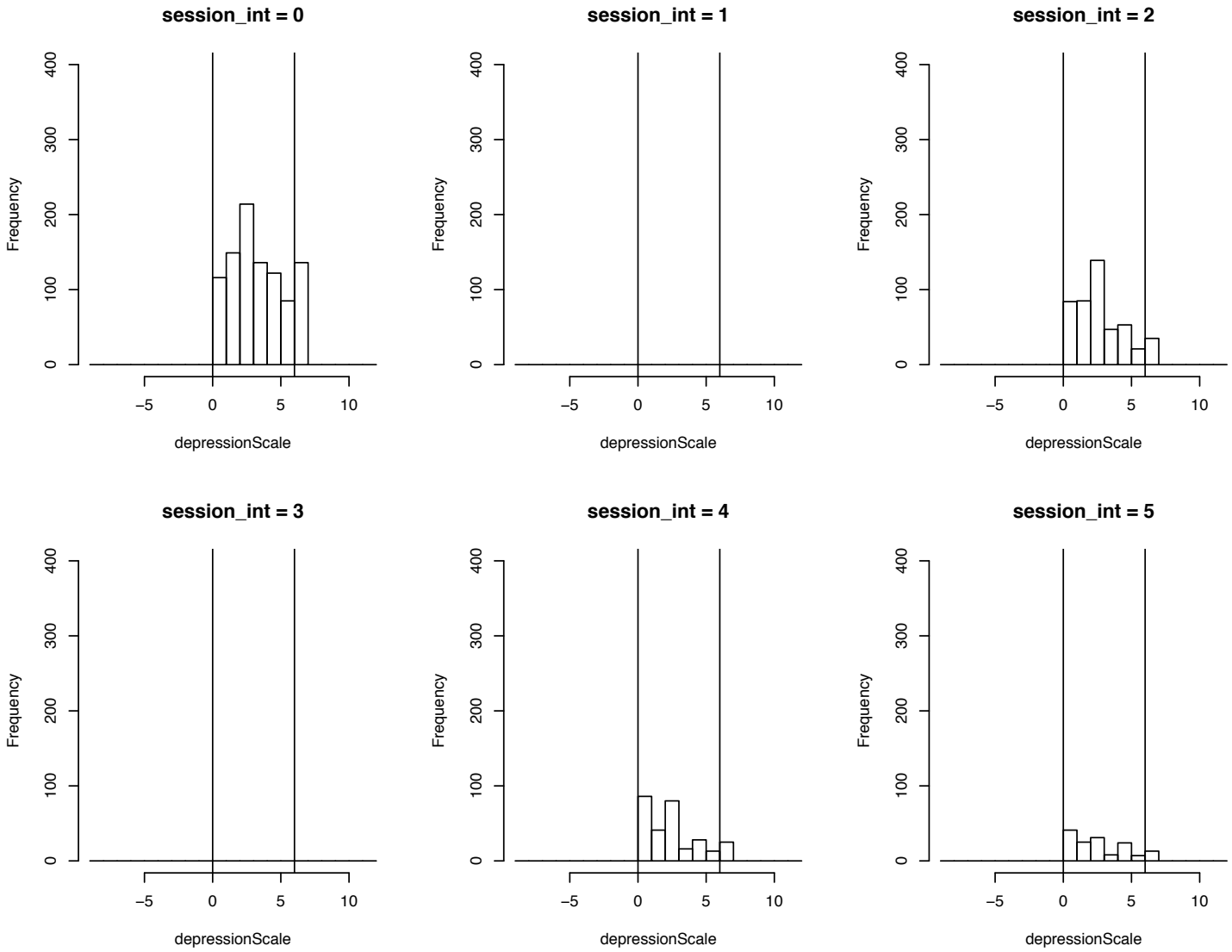

Figure S4.1. Distribution of depression scores at each time point before multiple imputation. Vertical lines show range of plausible values (cells are left-closed). Depression was not assessed at Sessions 1 or 3.

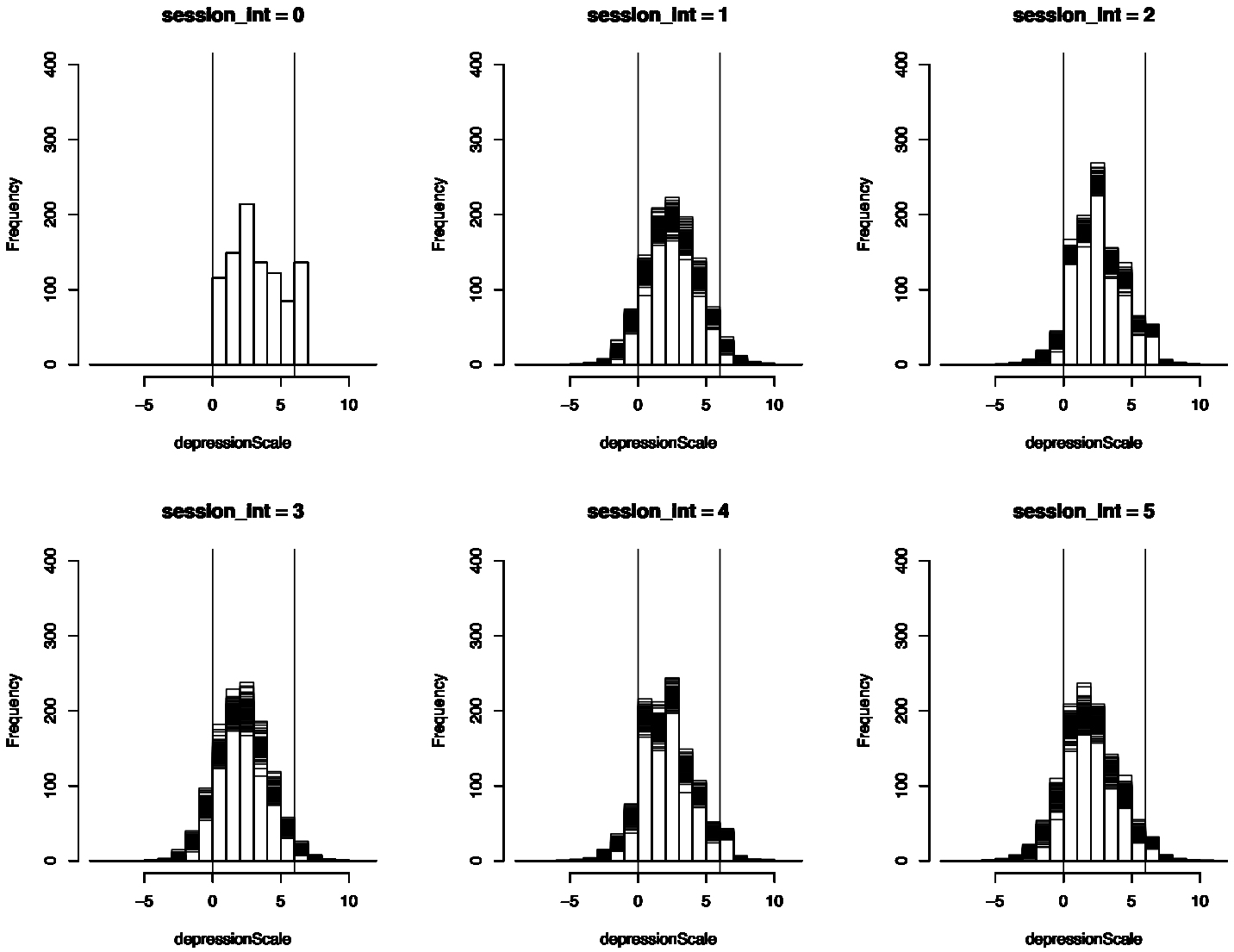

Figure S4.2. Overlaid distributions of depression scores at each time point for 100 multiple imputations of Combined-Level Dataset. Vertical lines show range of plausible values (cells are left-closed). Imputed scores for Sessions 1 and 3 were not analyzed.

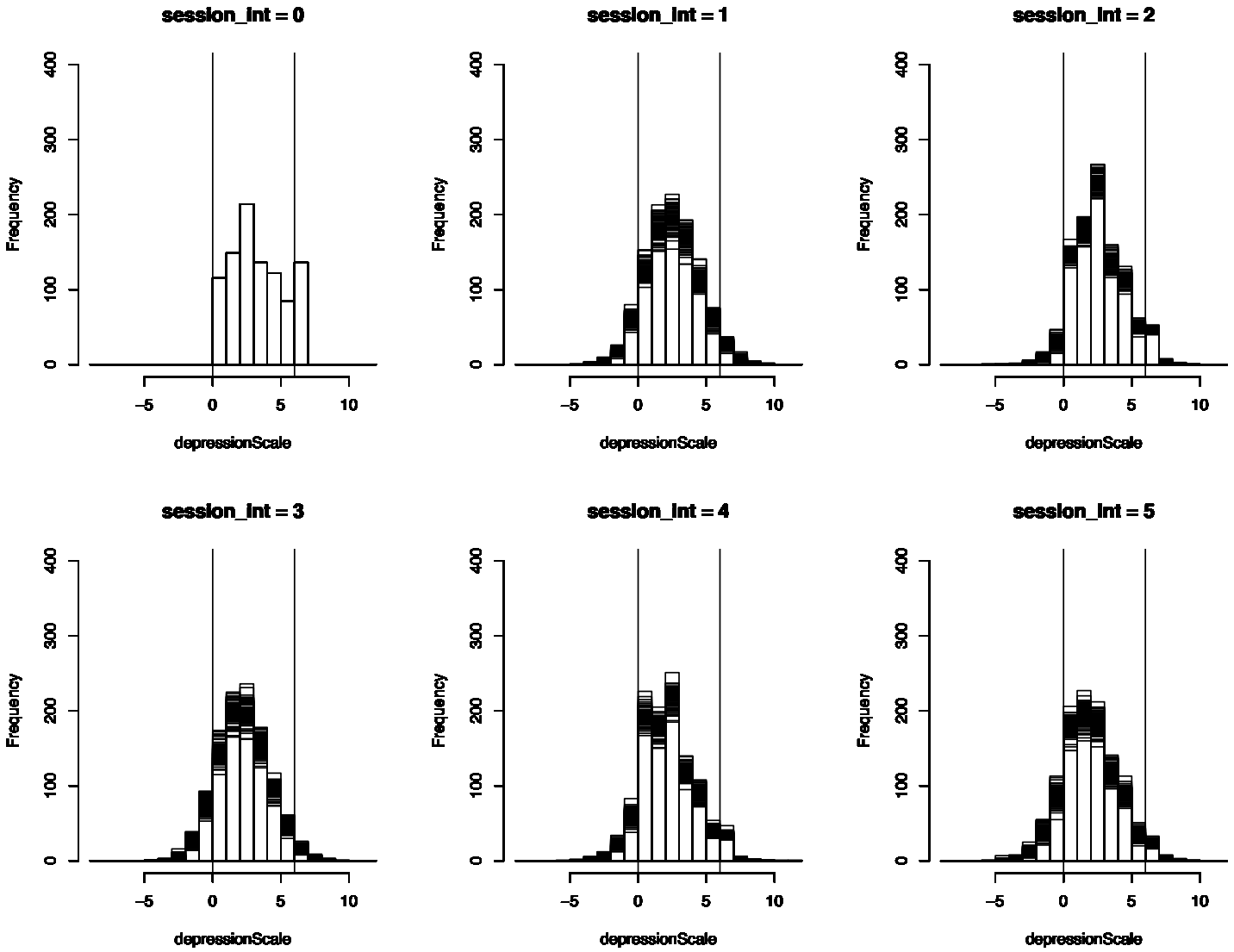

Figure S4.3. Overlaid distributions of depression scores at each time point for 100 multiple imputations of Separate-Level Dataset. Vertical lines show range of plausible values (cells are left-closed). Imputed scores for Sessions 1 and 3 were not analyzed.

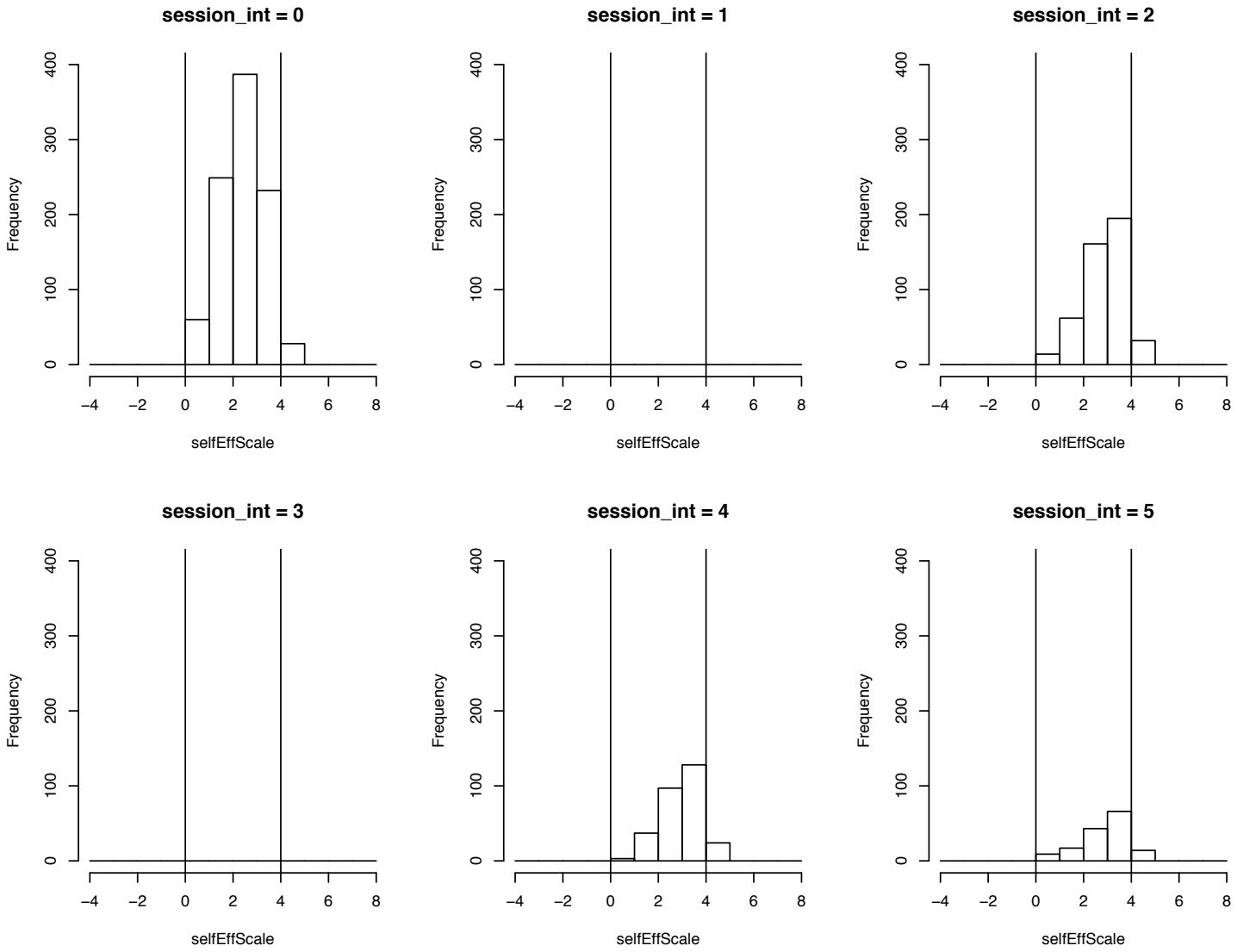

Figure S5.1. Distribution of self-efficacy scores at each time point before multiple imputation. Vertical lines show range of plausible values (cells are left-closed). Self-efficacy was not assessed at Sessions 1 or 3.

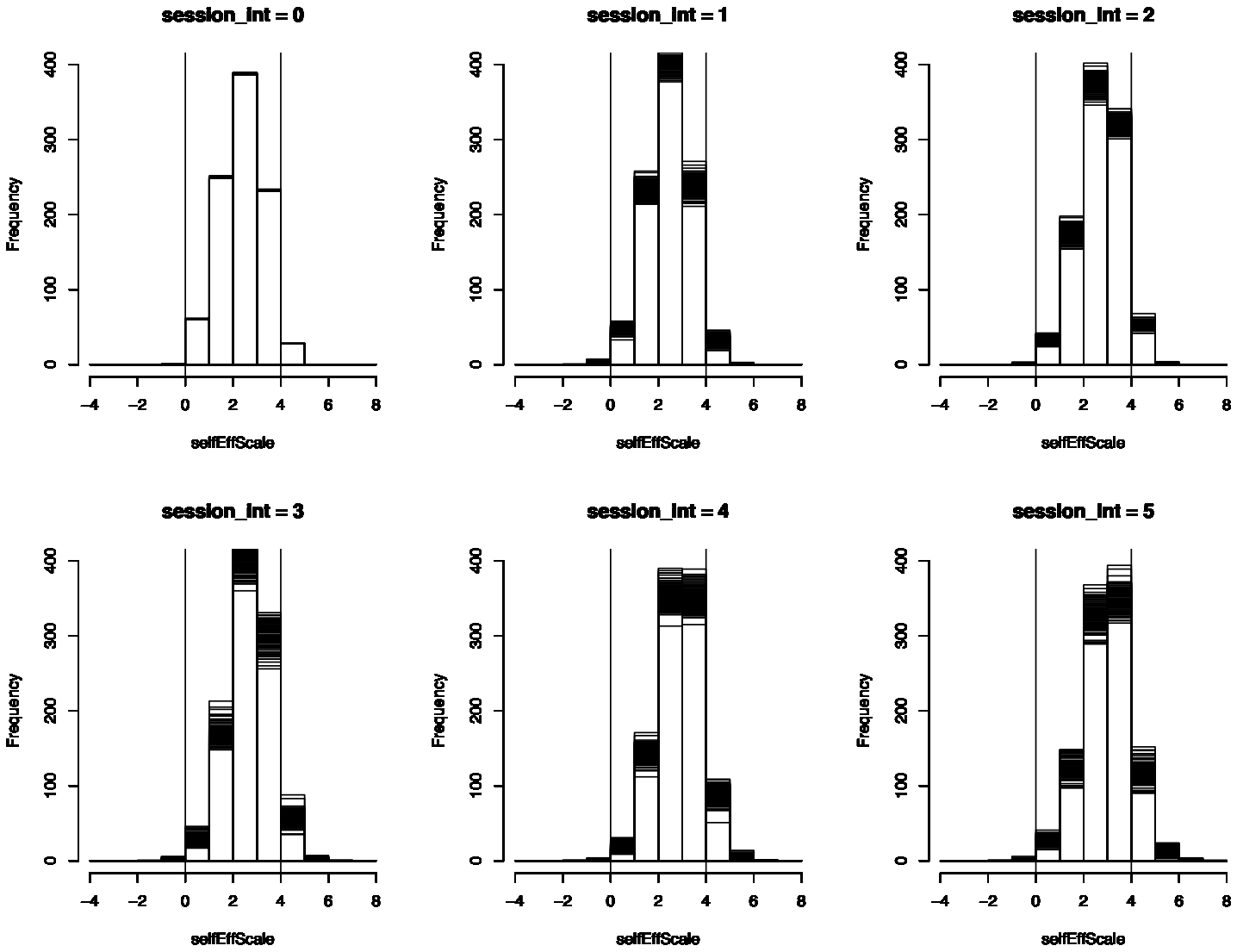

Figure S5.2. Overlaid distributions of self-efficacy scores at each time point for 100 multiple imputations of Combined-Level Dataset. Vertical lines show range of plausible values (cells are left-closed). Imputed scores for Sessions 1 and 3 were not analyzed.

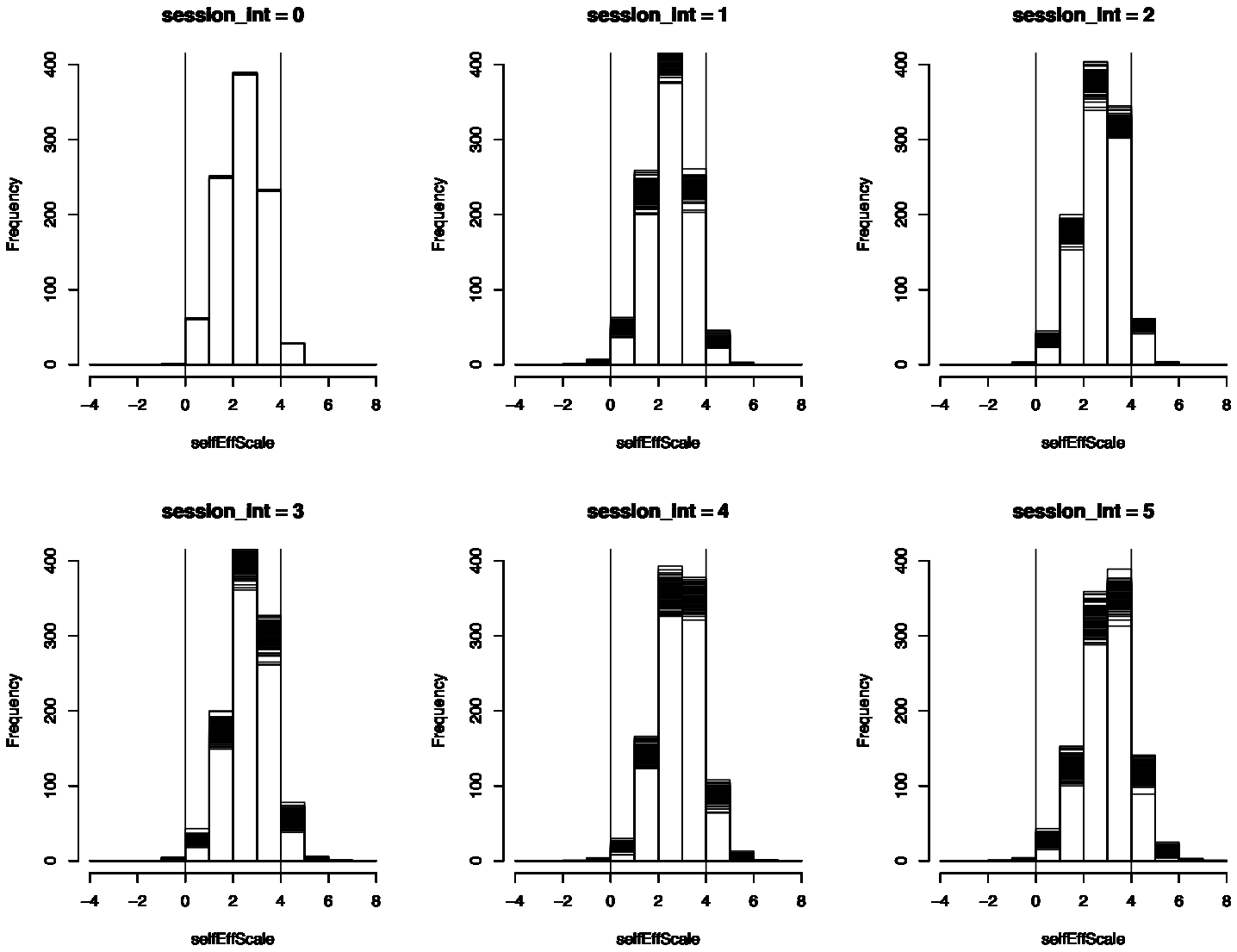

Figure S5.3. Overlaid distributions of self-efficacy scores at each time point for 100 multiple imputations of Separate-Level Dataset. Vertical lines show range of plausible values (cells are left-closed). Imputed scores for Sessions 1 and 3 were not analyzed.

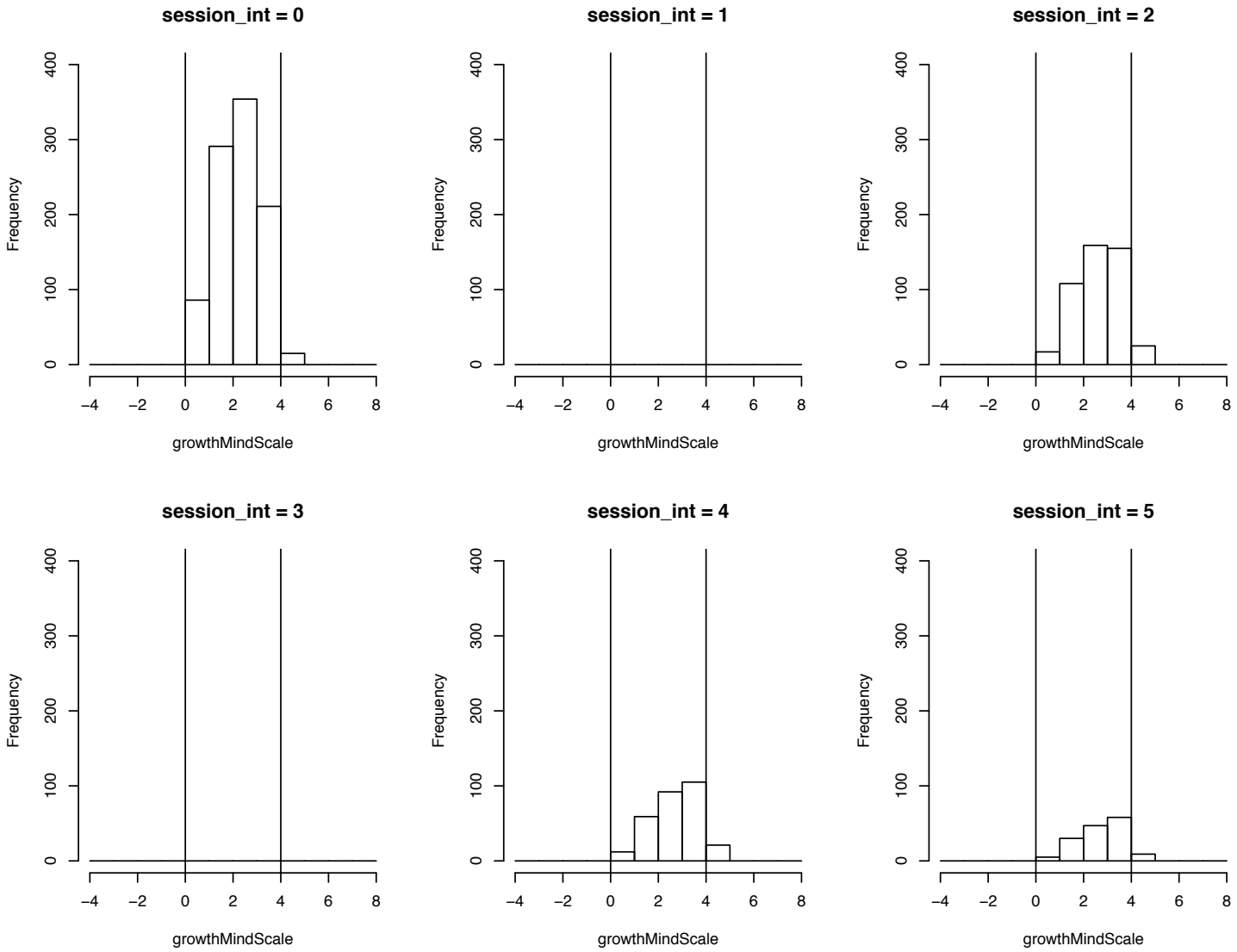

Figure S6.1. Distribution of growth mindset scores at each time point before multiple imputation. Vertical lines show range of plausible values (cells are left-closed). Growth mindset was not assessed at Sessions 1 or 3.

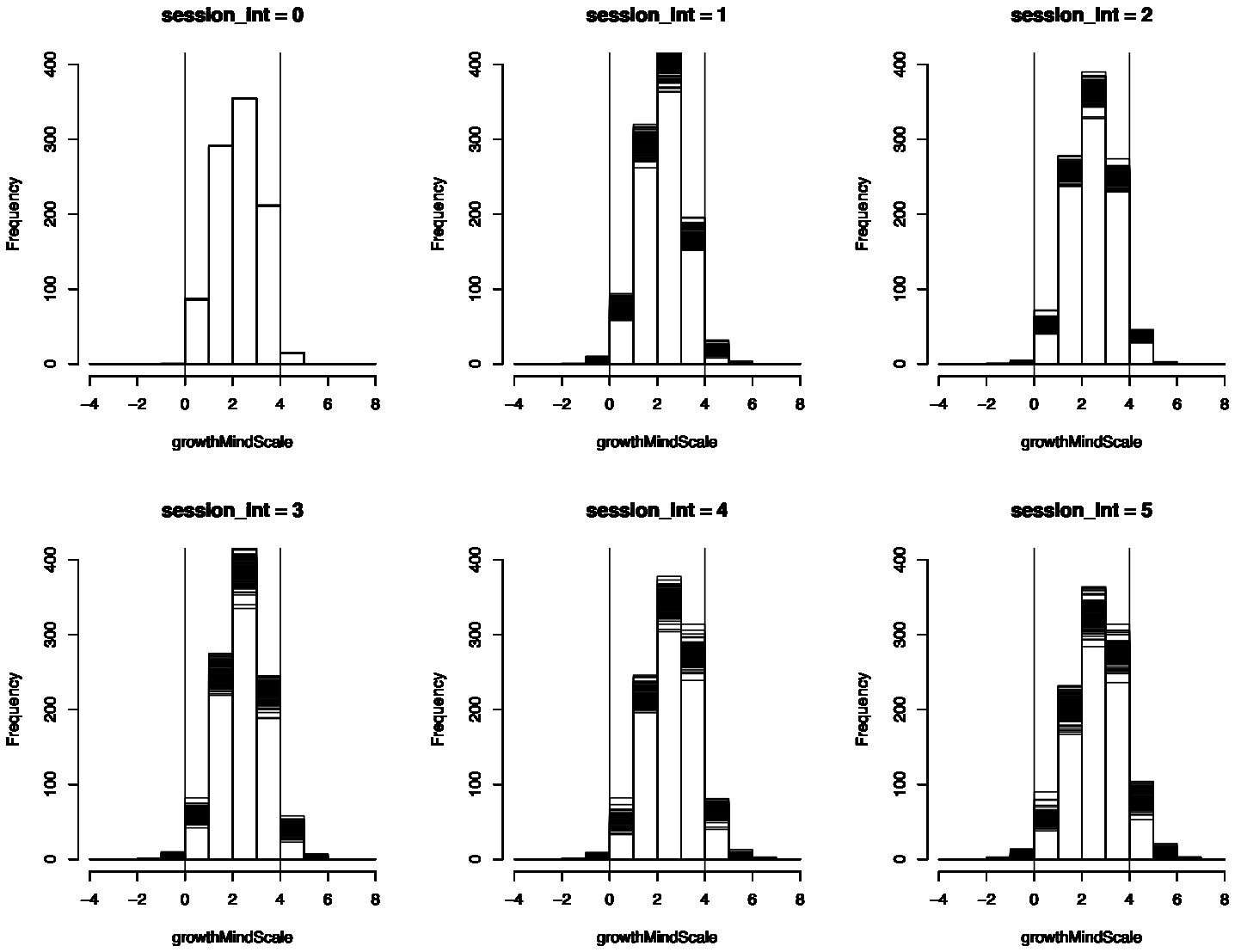

Figure S6.2. Overlaid distributions of growth mindset scores at each time point for 100 multiple imputations of Combined-Level Dataset. Vertical lines show range of plausible values (cells are left-closed). Imputed scores for Sessions 1 and 3 were not analyzed.

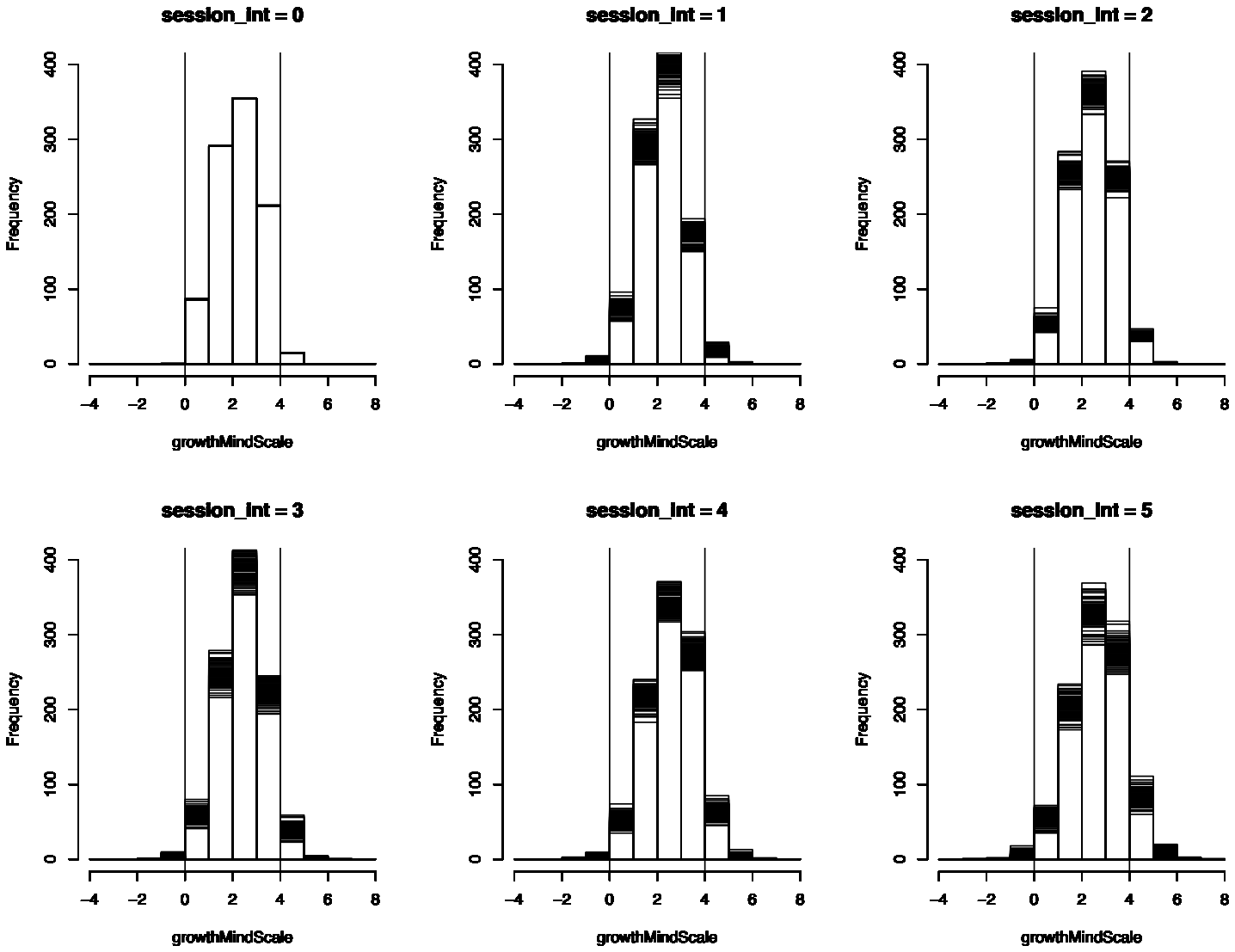

Figure S6.3. Overlaid distributions of growth mindset scores at each time point for 100 multiple imputations of Separate-Level Dataset. Vertical lines show range of plausible values (cells are left-closed). Imputed scores for Sessions 1 and 3 were not analyzed.

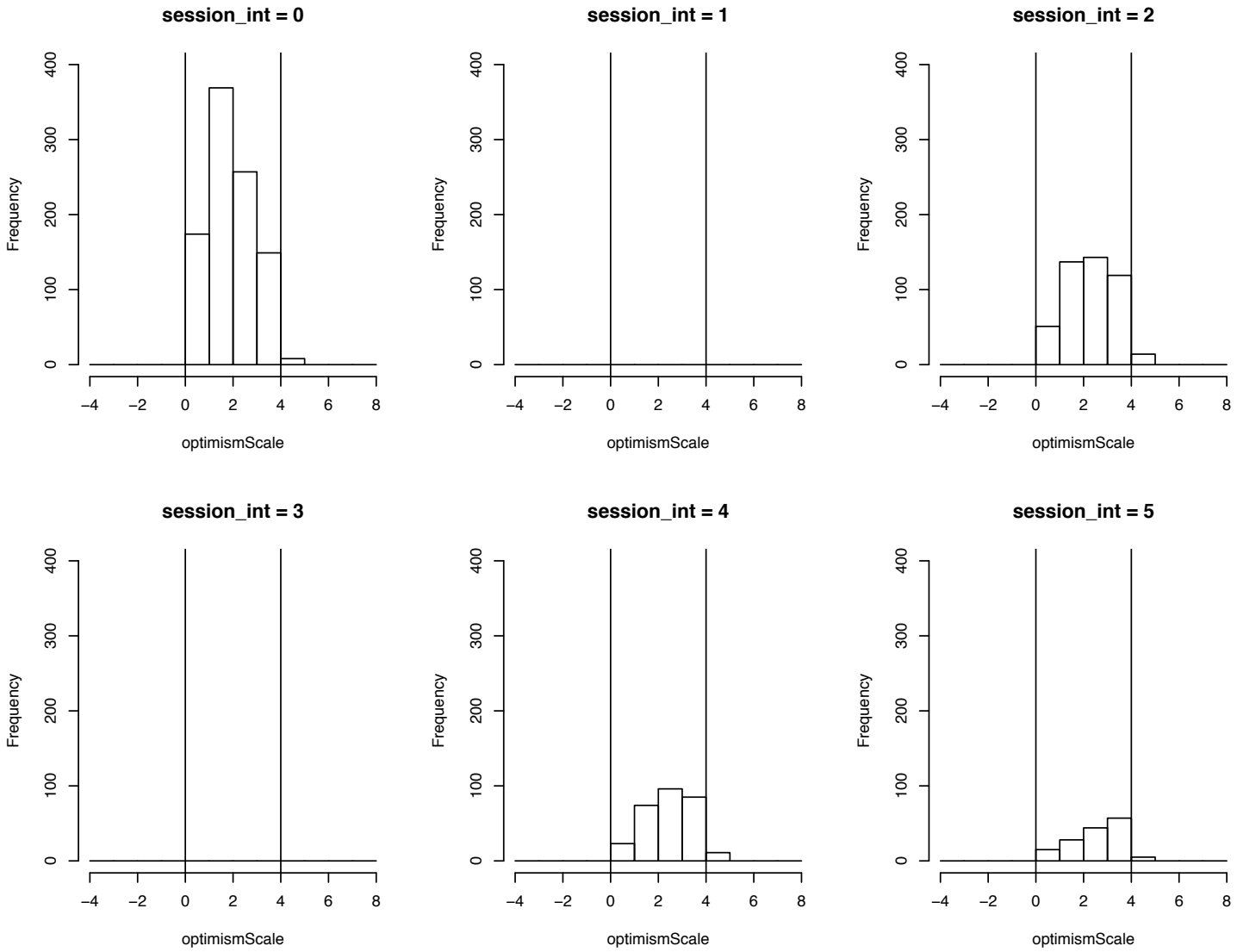

Figure S7.1. Distribution of optimism scores at each time point before multiple imputation. Vertical lines show range of plausible values (cells are left-closed). Optimism was not assessed at Sessions 1 or 3.

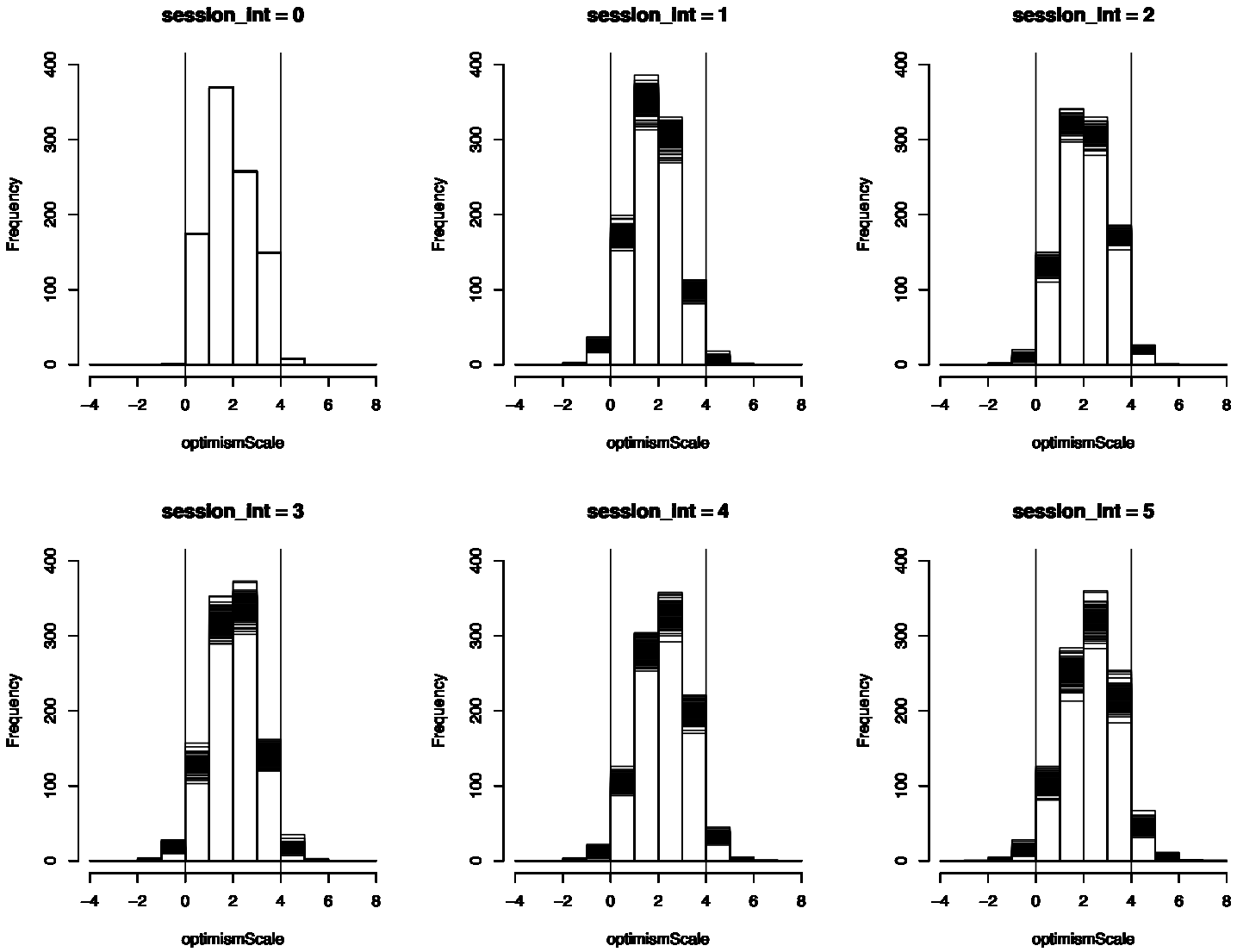

Figure S7.2. Overlaid distributions of optimism scores at each time point for 100 multiple imputations of Combined-Level Dataset. Vertical lines show range of plausible values (cells are left-closed). Imputed scores for Sessions 1 and 3 were not analyzed.

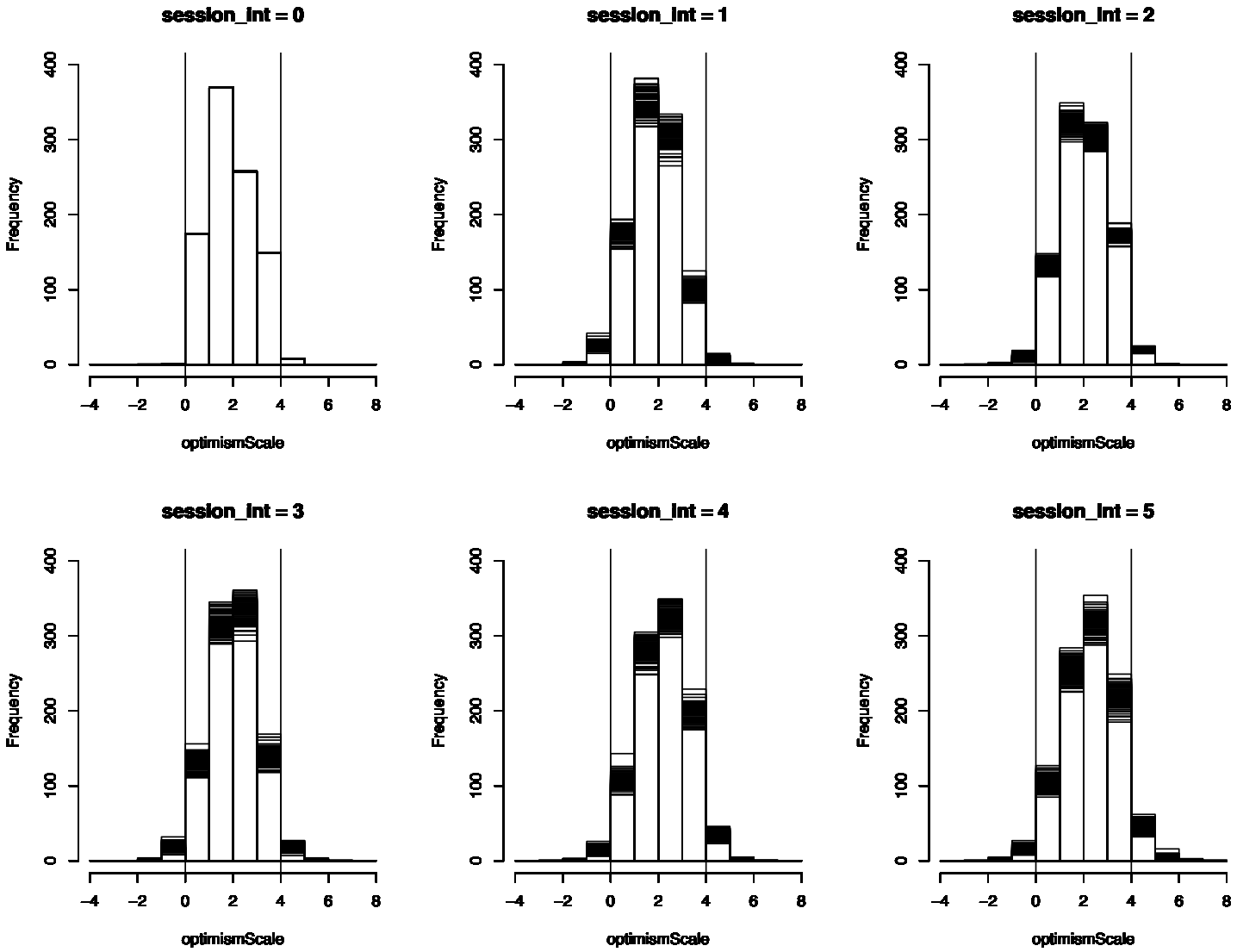

Figure S7.3. Overlaid distributions of optimism scores at each time point for 100 multiple imputations of Separate-Level Dataset. Vertical lines show range of plausible values (cells are left-closed). Imputed scores for Sessions 1 and 3 were not analyzed.