Supplemental Materials

Contents

Table S1. Participant Self-Reports of Impacts of the Intervention and Their Causes Based on the Psychotherapy Side Effects Scale (PSES) in Study 1.

Items	Causes (Participant_ID)
Item 1. Made me feel emotionally low	The review tasks had not been completed yet. (E9520)
	I felt somewhat frustrated, especially when encountering problems while performing the tasks. (E2579)
Item 2. Decreased my ability to work or study	It was a bit uncomfortable to be interrupted every 10 minutes. (E9520)
Item 3. Made me feel dependent on the treatment	I think the intervention process this time was gradual and well-paced, and I really liked that rhythm. (E1492)
	I may have realized that I can rely on external help to change myself. (E2398)
	The method was so good that I couldn't help wanting to use it again in the future. (T4693)
	The effect was so good. (E2825)
	The intervention could solve part of the problems I was worried about. (E0439)
	The approach and the way the conversation was initiated were excellent. (T4121)
	I believed the coach could relieve my anxiety and negative emotions, which made me want to rely on the coach. (T9889)
	I felt more confident about accomplishing certain things. (T8397)
Item 4. Made me feel anxious or uneasy	The intervention brought up the issue of the anxiety-inducing task again. (E9520)
	The intervention required me to confront anxiety-inducing tasks, and in the process I experienced feelings of anxiety and unease. (E1739)
	I was worried that I might not perform well (on my academic tasks) without the next intervention. (E2398)
Item 5. Made me feel pessimistic and hopeless	The difficulty of the task made me realize how hard it is to complete. (E9520)
	The anxiety was brought about by the task content. (E2579)

	It required me to confront anxiety-inducing tasks, during which I experienced feelings of anxiety and unease. (E1739)
Item 8. Made me more sensitive and suspicious	I was not sure whether this approach would be applicable to other tasks in the future. (E2398)
Item 12. Caused physical discomfort	Mild tension and numb bodily sensations associated with anxiety. (E2579)
Item 14. Strained my relationships with others	The intervention took a bit too long, and when I came out, my classmates didn't really understand why I was participating in it. (T4693)
Item 15. Made me overly confident or optimistic	Some of the positive evaluations might be somewhat exaggerated. (E5560)
	I realized that with external support, I could accomplish certain tasks. (E2398)
	I felt like I could solve every problem. (E2825)
	This intervention addressed some of my anxiety. (E1492)
Item 18. Made me feel ashamed	The intervention made me realize that I really hadn't been doing well enough before. (T4693)
Item 20. Failed to resolve my original problems	The task had not yet been completed, and the source of my anxiety still remained. (E9520)
	The duration of the intervention allowed only an initial improvement, and a longer period would be needed to fully benefit from it. (E0665)
Item 23. Made me feel emotionally overexcited or elated	My level of anxiety decreased, and my emotions also changed. (E1492)
	I gained some valuable insights. (E7201)
	I had gained confidence in my studies and felt more motivated. (T4693)
	I was very happy to complete this task. (E2825)
	Through the intervention, I realized that actively facing anxiety could alleviate it. (T7144)
	The encouraging comments could inspire me to become passionate. (E5560)
	I felt very happy, and it was a great help to me. (E0439)

Item 25. Made me act impulsively or recklessly	Completing the task made me a little more confident in myself. (E1492)

Note. The table presents English translations; original participant responses were provided in Chinese.

Table S2. Fixed Effects Estimates from the Linear Mixed-Effects Model Based on the Intention-to-Treat Sample in Study 1.

	B	SE	*p*
Primary Outcomes			
Academic anxiety			
Intercept	14.38	0.73	<.001
Group	0.40	0.47	.396
$Time_{baseline,\ follow}$	-1.22	0.49	.013
Group × $Time_{baseline,\ follow}$	-1.31	0.68	.059
Depressive symptoms at baseline	0.16	0.05	.001
Paternal education	0.09	0.19	.649
Procrastination			
Intercept	11.84	1.39	<.001
Group	1.83	0.92	.051
$Time_{baseline,\ follow}$	-0.68	0.59	.254
Group × $Time_{baseline,\ follow}$	-2.45	0.84	.004
Depressive symptoms at baseline	0.40	0.09	<.001
Paternal education	0.18	0.37	.630
State anxiety			
Intercept	38.77	2.99	<.001
Group	-0.67	2.10	.751
$Time_{baseline,\ post}$	-3.07	1.21	.013
$Time_{baseline,\ follow}$	-4.08	1.63	.014
Group × $Time_{baseline,\ post}$	-5.04	1.72	.004
Group × $Time_{baseline,\ follow}$	-1.32	2.30	.567
Depressive symptoms at baseline	1.20	0.20	<.001
Paternal education	-0.57	0.78	.463
Secondary Outcomes			
Positive emotion			
Intercept	28.55	2.21	<.001
Group	0.97	1.41	.493
$Time_{baseline,\ post}$	0.22	0.70	.751
$Time_{baseline,\ follow}$	1.09	0.91	.234
Group × $Time_{baseline,\ post}$	3.62	0.99	<.001
Group × $Time_{baseline,\ follow}$	1.62	1.29	.212
Depressive symptoms at baseline	-0.29	0.15	.053
Paternal education	0.48	0.59	.417
Negative emotion			
Intercept	14.79	1.85	<.001
Group	1.49	1.31	.256
$Time_{baseline,\ post}$	-2.53	0.86	.004
$Time_{baseline,\ follow}$	-1.71	0.93	.071
Group × $Time_{baseline,\ post}$	-3.89	1.22	.002

Group × $Time_{baseline, follow}$	-1.78	1.32	.179
Depressive symptoms at baseline	0.96	0.12	<.001
Paternal education	0.32	0.48	.508
Generalized anxiety			
Intercept	3.92	0.99	<.001
Group	0.49	0.73	.504
$Time_{baseline, follow}$	-2.16	0.51	<.001
Group × $Time_{baseline, follow}$	-0.82	0.71	.253
Depressive symptoms at baseline	0.55	0.07	<.001
Paternal education	-0.12	0.25	.638
Depressive symptoms			
Intercept	9.67	1.04	<.001
Group	-1.40	0.91	.127
$Time_{baseline, follow}$	-1.66	0.48	.001
Group × $Time_{baseline, follow}$	-0.03	0.68	.964
Paternal education	-0.76	0.36	.040
Mechanism Variables			
Distress tolerance			
Intercept	22.13	1.27	<.001
Group	0.37	0.87	.674
$Time_{baseline, post}$	0.69	0.44	.122
$Time_{baseline, follow}$	0.30	0.58	.608
Group × $Time_{baseline, post}$	0.98	0.62	.120
Group × $Time_{baseline, follow}$	0.97	0.82	.244
Depressive symptoms at baseline	-0.17	0.08	.042
Paternal education	0.19	0.34	.583
Self-efficacy			
Intercept	7.39	0.42	<.001
Group	0.01	0.29	.972
$Time_{baseline, post}$	0.24	0.17	.155
$Time_{baseline, follow}$	0.22	0.18	.233
Group × $Time_{baseline, post}$	0.93	0.24	<.001
Group × $Time_{baseline, follow}$	0.73	0.26	.006
Depressive symptoms at baseline	-0.08	0.03	.003
Paternal education	-0.18	0.11	.112

Note. The Group variable was coded with the control group as the reference category. The Time variable was coded with baseline as the reference point.

Table S3. Estimated Means of Variables at Each Assessment Point Based on Linear Mixed-Effects Models and the Corresponding Within- and Between-Group Change Statistics for the Complete-Case Sample in Study 1 (N = 89).

Variables	Intervention Group (N = 45)			Control Group (N = 44)			Between-Group Differences in Change	
	M (SE)	Within-group Change		M (SE)	Within-group Change			
		d (95% CI)	*p* value		*d* (95% CI)	*p* value	*d* (95% CI)	*p* value
Primary Outcomes								
Academic anxiety								
Baseline	16.16 (0.33)			15.77 (0.33)				
Follow up	13.63 (0.43)	1.18 [0.74, 1.61]	<.001	14.50 (0.44)	0.53 [0.13, 0.93]	.010	0.55 [-0.03, 1.14]	.068
Procrastination								
Baseline	16.94 (0.61)			15.20 (0.62)				
Follow up	13.81 (0.58)	0.70 [0.45, 0.96]	<.001	14.38 (0.59)	0.18 [-0.08, 0.44]	.168	0.52 [0.16, 0.89]	.006
State anxiety								
Baseline	44.89 (1.44)			45.64 (1.46)				
Postintervention	36.77 (1.39)	0.82 [0.58, 1.06]	<.001	42.44 (1.41)	0.24 [0.06, 0.43]	.022	0.42 [0.13, 0.71]	.006
Follow up	39.49 (1.48)	0.55 [0.23, 0.87]	.003	41.39 (1.50)	0.32 [0.08, 0.57]	.022	0.10 [-0.29, 0.49]	.618
Secondary Outcomes								
Positive emotion								
Baseline	28.83 (0.98)			27.86 (0.99)				
Postintervention	32.67 (0.99)	0.61 [0.39, 0.83]	<.001	28.09 (1.00)	0.03 [-0.17, 0.24]	.750	0.55 [0.25, 0.85]	.001
Follow up	31.54 (1.03)	0.43 [0.15, 0.72]	.012	28.93 (1.04)	0.16 [-0.11, 0.42]	.496	0.25 [-0.14, 0.63]	.207
Negative emotion								
Baseline	23.69 (0.91)			22.27 (0.92)				
Postintervention	17.27 (0.86)	0.92 [0.68, 1.17]	<.001	19.61 (0.87)	0.33 [0.12, 0.55]	.006	0.50 [0.18, 0.82]	.003
Follow up	20.20 (0.90)	0.50 [0.24, 0.76]	<.001	20.47 (0.91)	0.23 [-0.01, 0.46]	.058	0.23 [-0.12, 0.57]	.202

Generalized anxiety								
Baseline	7.85 (0.50)			7.41 (0.51)				
Follow up	4.87 (0.40)	0.69 [0.47, 0.92]	<.001	5.18 (0.41)	0.51 [0.28, 0.74]	<.001	0.17 [-0.15, 0.49]	.296
Depressive symptoms								
Baseline	6.28 (0.56)			7.35 (0.57)				
Follow up	4.59 (0.54)	0.44 [0.20, 0.67]	<.001	5.85 (0.55)	0.32 [0.12, 0.51]	.002	0.04 [-0.26, 0.35]	.776
Mechanism Variables								
Distress tolerance								
Baseline	21.70 (0.59)			21.30 (0.59)				
Postintervention	23.37 (0.51)	0.43 [0.21, 0.65]	<.001	22.03 (0.51)	0.18 [-0.04, 0.41]	.216	0.24 [-0.07, 0.55]	.139
Follow up	22.97 (0.62)	0.32 [0.03, 0.62]	.093	21.67 (0.62)	0.09 [-0.20, 0.38]	.536	0.23 [-0.18, 0.64]	.276
Self-efficacy								
Baseline	6.35 (0.20)			6.34 (0.20)				
Postintervention	7.53 (0.18)	0.79 [0.56, 1.01]	<.001	6.59 (0.18)	0.19 [-0.06, 0.45]	.304	0.66 [0.33, 0.99]	<.001
Follow up	7.31 (0.21)	0.64 [0.40, 0.88]	<.001	6.57 (0.22)	0.18 [-0.11, 0.46]	.304	0.52 [0.16, 0.88]	.006

Note. Within-group change p-values were adjusted using the Holm–Bonferroni correction within each variable to control for multiple comparisons.

Table S4. Fixed-Effects Estimates from the Linear Mixed-Effects Model for the Complete-Case Sample in Study 1 (N = 89).

	B	SE	*p*
Primary Outcomes			
Academic anxiety			
Intercept	14.21	0.73	<.001
Group	0.39	0.47	.409
$Time_{baseline,\ follow}$	-1.27	0.49	.010
Group × $Time_{baseline,\ follow}$	-1.26	0.68	.068
Depressive symptoms at baseline	0.20	0.05	<.001
Paternal education	0.08	0.19	.683
Procrastination			
Intercept	11.16	1.34	<.001
Group	1.74	0.88	.051
$Time_{baseline,\ follow}$	-0.82	0.59	.168
Group × $Time_{baseline,\ follow}$	-2.32	0.83	.006
Depressive symptoms at baseline	0.53	0.10	<.001
Paternal education	0.16	0.35	.642
State anxiety			
Intercept	38.25	3.02	<.001
Group	-0.75	2.07	.717
$Time_{baseline,\ post}$	-3.20	1.23	.011
$Time_{baseline,\ follow}$	-4.25	1.63	.011
Group × $Time_{baseline,\ post}$	-4.91	1.73	.006
Group × $Time_{baseline,\ follow}$	-1.15	2.30	.618
Depressive symptoms at baseline	1.32	0.22	<.001
Paternal education	-0.61	0.78	.435
Secondary Outcomes			
Positive emotion			
Intercept	28.09	2.25	<.001
Group	0.97	1.41	.496
$Time_{baseline,\ post}$	0.23	0.71	.750
$Time_{baseline,\ follow}$	1.07	0.92	.248
Group × $Time_{baseline,\ post}$	3.62	1.00	.001
Group × $Time_{baseline,\ follow}$	1.64	1.29	.207
Depressive symptoms at baseline	-0.21	0.17	.204
Paternal education	0.47	0.59	.427
Negative emotion			
Intercept	14.63	1.88	<.001
Group	1.42	1.30	.277
$Time_{baseline,\ post}$	-2.66	0.87	.003
$Time_{baseline,\ follow}$	-1.80	0.94	.058
Group × $Time_{baseline,\ post}$	-3.76	1.23	.003

Group × $Time_{baseline, follow}$	-1.69	1.32	.202
Depressive symptoms at baseline	1.01	0.14	<.001
Paternal education	0.30	0.48	.531
Generalized anxiety			
Intercept	3.72	0.98	<.001
Group	0.44	0.72	.539
$Time_{baseline, follow}$	-2.23	0.51	<.001
Group × $Time_{baseline, follow}$	-0.75	0.71	.296
Depressive symptoms at baseline	0.59	0.07	<.001
Paternal education	-0.12	0.25	.629
Depressive symptoms			
Intercept	8.87	0.96	<.001
Group	-1.07	0.81	.190
$Time_{baseline, follow}$	-1.50	0.47	.002
Group × $Time_{baseline, follow}$	-0.19	0.66	.776
Paternal education	-0.59	0.33	.080
Mechanism Variables			
Distress tolerance			
Intercept	22.78	1.25	<.001
Group	0.40	0.84	.637
$Time_{baseline, post}$	0.73	0.45	.108
$Time_{baseline, follow}$	0.36	0.59	.536
Group × $Time_{baseline, post}$	0.94	0.63	.139
Group × $Time_{baseline, follow}$	0.90	0.82	.276
Depressive symptoms at baseline	-0.30	0.09	.002
Paternal education	0.21	0.32	.517
Self-efficacy			
Intercept	7.47	0.43	<.001
Group	0.01	0.29	.972
$Time_{baseline, post}$	0.25	0.17	.152
$Time_{baseline, follow}$	0.23	0.19	.224
Group × $Time_{baseline, post}$	0.93	0.24	<.001
Group × $Time_{baseline, follow}$	0.73	0.26	.006
Depressive symptoms at baseline	-0.10	0.03	.002
Paternal education	-0.17	0.11	.123

Note. The Group variable was coded with the control group as the reference category. The Time variable was coded with baseline as the reference point

Table S5. Feedback from the participant who scored 37 on the Psychotherapy Side Effects Scale (PSES) in Study 2.

Items	Reasons
Item 4. Made me feel anxious or uneasy (Mild)	Because the task itself was anxiety-provoking, I felt uneasy.
Item 6. Made me want to stop the treatment (Mild)	I felt that chatting with the chatbot took too much time.
Item 7. Caused negative emotions (Moderate)	The negative feelings arose because I realized I was very slow at memorizing words and kept forgetting them.
Item 8. Made me more sensitive and suspicious (Mild)	I began to doubt my own memory and learning ability, which made me feel more sensitive and suspicious.
Item 9. Worsened my existing problems (Mild)	I was worried that I might forget the idea of facing and coexisting with anxiety in the next session and wouldn't be able to start again.
Item 13. Triggered feelings of anger (Mild)	The anger was probably because I couldn't remember the words I had learned before.
Item 15. Made me overly confident or optimistic (Moderate)	The chatbot kept encouraging me and never pointed out my mistakes, even though I knew my performance was poor. I felt that if it were a real teacher, they might have been annoyed.
Item 18. Made me feel ashamed (Mild)	The shame came from repeatedly talking about my weaknesses, such as my poor English and failure to pass CET-6.

Item 20. Failed to resolve my original problems (Mild)	The issue of anxiety can’t be solved in one session — it takes multiple attempts.
Item 24. Made me feel inferior (Mild)	The reason for feeling inferior was the same as for feeling ashamed.

Note. The table presents English translations; original participant responses were provided in Chinese.

Table S6. Participant Self-Reports of Impacts of the Intervention and Their Causes Based on the Psychotherapy Side Effects Scale (PSES) in Study 2.

Items	Causes (Participant_ID)
Item 1. Made me feel emotionally low	I had to face the things that make me anxious directly, and the whole process just made me feel really down. (SSE7301, SSE2383)
	The task I did didn't really achieve the results I was looking for. (SSE1810, SSE9214)
	The practice task was just too hard for me. (SSE9379)
	The task made me realize how tough the job market is, which really brought me down and made me anxious (SSE7049)
Item 2. Decreased my ability to work or study	Having an experimenter watching me while I did the tasks made me feel a bit uneasy. (SSE8581)
	Stopping every ten minutes to talk to the chatbot made it hard for me to focus on the task. (SSE1618)
Item 3. Made me feel dependent on the treatment	I felt the whole treatment was very helpful, and I want to keep doing it. (SSE0135, SSE2795, SSE4524, SSE6368, SSE7675, SSE6856, SSE5256, SSE6646, SSE3497, SSE7788, SSE6468, SSE3565)
	The intervention helped me reduce my anxiety. I want to do it again the next time I feel anxious. (SSE7014, SSE7155, SSE9683, SSE8187, SSE1976, SSE5784, SSE1686, SSE4142, SSE3183)
	During the intervention, the coach and the chatbot let me talk about my feelings and gave me encouragement. I felt an emotional connection, and it felt really good. (SSE2465, SSE6718, SSE1188, SSE2875, SSE7301, SSE7218)
	It was easier for me to focus on my schoolwork during the exposure exercise because someone was watching me. (SSE1856, SSE5307, SSE1489, SSE5530, SSE7883, SSE2383)

Item 4. Made me feel anxious or uneasy	I had to face the things that made me anxious, and it brought back my feelings of anxiety. (SSE1489, SSE6856, SSE3183, SSE7282, SSE8190, SSE5783, SSE7281, SSE3536)
	The practice tasks were too hard, and things didn't go well. (SSE2875, SSE9379, SSE4142)
	I was a bit nervous at the beginning. I was afraid to say what I was thinking. (SSE3981)
	I knew that I was being watched and monitored. (SSE7991)
	I was interrupted every ten minutes, so it was hard for me to get into the task. (SSE6208)
Item 6. Made me want to stop the treatment	During the exposure exercise, I didn't want to spend so much time chatting with the robot every ten minutes. I wanted to just keep working on my task. (SSE6208)
	It was a bit painful to talk about my anxious feelings and thoughts over and over again. (SSE2383)
Item 7. Caused negative emotions	I had to do tasks that made me feel anxious. (SSE0409, SSE9214, SSE3183, SSE7281, SSE7282, SSE2383)
	The tasks didn't go well. (SSE2875)
	The 10-minute countdown on the webpage made me feel anxious. (SSE6004)
Item 8. Made me more sensitive and suspicious	The intervention made me remember bad experiences from the past, and I started to doubt myself. (SSE0409)
	The intervention made me feel that things wouldn't always happen the way I imagined. It made me doubt my own thoughts. (SSE7301)

Item 9. Worsened my existing problems	I ran into some new problems during the exposure exercise. (SSE2875)
Item 11. Made me feel that someone was trying to harm me	The result of my task was not satisfactory (SSE9214)
Item 12. Caused physical discomfort	The anxiety from the exposure tasks made me feel uncomfortable. (SSE7281, SSE2383)
	Having headaches due to long time learning (SSE7282)
Item 13. Triggered feelings of anger	I didn't do well during the exposure exercise, and I felt angry with myself. (SSE3284, SSE3183, SSE1163)
Item 14. Strained my relationships with others	The intervention made me think about anxious things, so I felt upset and didn't want to talk to anyone. (SSE6004)
Item 15. Made me overly confident or optimistic	Because the tasks went well, I felt positive and confident. But I thought this confidence might only be temporary, and I would go back to reality later. (SSE0505, SSE6290, SSE1188, SSE1976, SSE1686, SSE8190, SSE3497, SSE5256, SSE1810, SSE3183, SSE7788, SSE3981, SSE2875, SSE3565, SSE7301, SSE2383)
	I felt the chatbot's encouragement was fake. I didn't actually do very well, and I thought it would just give me blind confidence. (SSE6905)
Item 17. Caused unrealistic romantic feelings toward the therapist	The therapist is polite and emotionally stable (SSE9033)
	The therapist is supportive and insightful (SSE2087, SSE4142, SSE7788, SSE3981)

	The therapist looks handsome (SSE6468)
Item 18. Made me feel ashamed	Reflecting on my problems and talking about my thoughts all the time made me feel ashamed. (SSE5162)
	It was awkward to have someone watching me work. (SSE3136)
	I had to talk about my secret thoughts, and it didn't feel good to talk about them. (SSE2383)
Item 20. Failed to resolve my original problems	This intervention wasn't very helpful to me. (SSE8581, SSE6208, SSE4142, SSE1810, SSE6905, SSE7281, SSE7282)
	The intervention is effective. However, a single session can hardly bring about a complete change and multiple interventions are required. (SSE5256)
Item 22. Made me hear people talking about me	I felt like someone was watching me, but I'm not really sure. (SSE5326)
Item 23. Made me feel emotionally overexcited or elated	The intervention was very effective and the task is progressing smoothly. I feel more confident and have learned a lot. Knowing that there is a solution to the problem makes me feel both happy and excited. (SSE0049, SSE2191, SSE7346, SSE8123, SSE4395, SSE5530, SSE7883, SSE8190, SSE3497, SSE7876, SSE7788, SSE3981, SSE3565, SSE2383)
	When I felt anxious about not being able to memorize words, unpleasant memories would come to mind and I would become agitated. (SSE3284)
	I'm feeling a bit nervous. (SSE7301)
Item 24. Made me feel inferior	I struggle to reach my ideal state, which leads to feelings of inferiority and self-doubt. (SSE2087, SSE7281, SSE2383)
	I feel like chatting with a chatbot is just a way for me to escape reality. (SSE6905)

Item 25. Made me act impulsively or recklessly	I opened up too much and shared too many of my real thoughts. Now I feel like I shouldn't have. (SSE1810)
	I was slightly affected by the environment (SSE3565)
	I was a bit nervous about this intervention and came in feeling unprepared, which led to things not going smoothly. (SSE6905)

Note. Responses were originally collected in Chinese and are presented here in English translation. Conceptually similar responses were grouped and summarized for clarity.

Table S7. Fixed Effects Estimates from the Linear Mixed-Effects Model Based on the Intention-to-Treat Sample in Study 2

	B	SE	*p*
Primary Outcomes			
Academic anxiety			
Intercept	12.02	0.45	<.001
Group	0.56	0.28	.048
$Time_{baseline,\ follow}$	-2.80	0.29	<.001
Group × $Time_{baseline,\ follow}$	-0.94	0.41	.023
Negative emotion at baseline	0.12	0.02	<.001
Procrastination			
Intercept	15.45	0.70	<.001
Group	0.94	0.47	.046
$Time_{baseline,\ follow}$	-2.88	0.40	<.001
Group × $Time_{baseline,\ follow}$	-1.40	0.57	.015
Negative emotion at baseline	0.10	0.02	<.001
State anxiety			
Intercept	29.24	1.32	<.001
Group	0.72	0.90	.426
$Time_{baseline,\ post}$	-8.95	0.81	<.001
$Time_{baseline,\ follow}$	-6.38	0.92	<.001
Group × $Time_{baseline,\ post}$	-2.81	1.16	.016
Group × $Time_{baseline,\ follow}$	-2.19	1.30	.094
Negative emotion at baseline	0.77	0.05	<.001
Secondary Outcomes			
Positive emotion			
Intercept	30.43	1.14	<.001
Group	-0.18	0.73	.805
$Time_{baseline,\ post}$	2.27	0.50	<.001
$Time_{baseline,\ follow}$	2.08	0.52	<.001
Group × $Time_{baseline,\ post}$	0.87	0.72	.226
Group × $Time_{baseline,\ follow}$	0.28	0.73	.699
Negative emotion at baseline	-0.07	0.04	.065
Negative emotion			
Intercept	25.35	0.63	<.001
Group	-1.98	0.89	.027
$Time_{baseline,\ post}$	-5.90	0.54	<.001
$Time_{baseline,\ follow}$	-3.87	0.63	<.001
Group × $Time_{baseline,\ post}$	-1.11	0.77	.148
Group × $Time_{baseline,\ follow}$	-0.70	0.89	.430
Generalized anxiety			
Intercept	2.18	0.63	.001
Group	0.14	0.40	.728

$Time_{baseline,\ follow}$	-2.95	0.33	<.001
Group × $Time_{baseline,\ follow}$	0.07	0.47	.875
Negative emotion at baseline	0.28	0.02	<.001
Depressive symptoms			
Intercept	4.08	0.72	<.001
Group	-0.43	0.46	.354
$Time_{baseline,\ follow}$	-2.60	0.33	<.001
Group × $Time_{baseline,\ follow}$	-0.51	0.47	.271
Negative emotion at baseline	0.21	0.03	<.001
Mechanism Variables			
Distress tolerance			
Intercept	21.84	0.69	<.001
Group	0.15	0.49	.766
$Time_{baseline,\ post}$	1.93	0.27	<.001
$Time_{baseline,\ follow}$	1.19	0.35	.001
Group × $Time_{baseline,\ post}$	0.70	0.39	.077
Group × $Time_{baseline,\ follow}$	0.50	0.49	.307
Negative emotion at baseline	-0.02	0.02	.465
Self-efficacy			
Intercept	6.73	0.25	<.001
Group	0.00	0.18	.982
$Time_{baseline,\ post}$	1.18	0.11	<.001
$Time_{baseline,\ follow}$	0.82	0.14	<.001
Group × $Time_{baseline,\ post}$	0.21	0.16	.206
Group × $Time_{baseline,\ follow}$	0.39	0.19	.042
Negative emotion at baseline	-0.04	0.01	<.001

Note. The Group variable was coded with the control group as the reference category. The Time variable was coded with baseline as the reference point.

Table S8. Estimated Means of Variables at Each Assessment Point Based on Linear Mixed-Effects Models and the Corresponding Within- and Between-Group Change Statistics for the Complete-Case Sample in Study 2. (N = 294)

Variables	Intervention Group (N = 140)			Control Group (N = 154)			Between-Group Differences in Change	
	M (SE)	Within-group Change		M (SE)	Within-group Change			
		d (95% CI)	*p* value		*d* (95% CI)	*p* value	*d* (95% CI)	*p* value
Primary Outcomes								
Academic Anxiety								
Baseline	15.47 (0.21)			15.00 (0.20)				
Follow up	11.65 (0.29)	1.63 [1.37, 1.88]	<.001	12.18 (0.28)	0.97 [0.77, 1.16]	<.001	0.37 [0.07, 0.68]	.019
Procrastination								
Baseline	19.00 (0.35)			18.09 (0.34)				
Follow up	14.55 (0.38)	1.06 [0.86, 1.26]	<.001	15.13 (0.36)	0.70 [0.51, 0.89]	<.001	0.35 [0.08, 0.63]	.012
State Anxiety								
Baseline	48.08 (0.67)			48.08 (0.64)				
Postintervention	36.66 (0.74)	1.10 [0.93, 1.26]	<.001	38.97 (0.70)	0.75 [0.61, 0.88]	<.001	0.20 [0.00, 0.41]	.057
Follow up	39.90 (0.90)	0.79 [0.60, 0.97]	<.001	41.56 (0.86)	0.53 [0.39, 0.68]	<.001	0.15 [-0.08, 0.38]	.217
Secondary Outcomes								
Positive emotion								
Baseline	28.63 (0.55)			28.36 (0.53)				
Postintervention	31.62 (0.59)	0.50 [0.33, 0.68]	<.001	30.63 (0.56)	0.32 [0.18, 0.46]	<.001	0.11 [-0.11, 0.33]	.325
Follow up	30.91 (0.59)	0.38 [0.20, 0.57]	<.001	30.49 (0.57)	0.30 [0.16, 0.44]	<.001	0.02 [-0.20, 0.25]	.837
Negative emotion								
Baseline	23.16 (0.69)			25.44 (0.66)				
Postintervention	16.40 (0.53)	0.90 [0.75, 1.06]	<.001	19.45 (0.51)	0.69 [0.56, 0.81]	<.001	0.10 [-0.10, 0.29]	.337
Follow up	18.57 (0.59)	0.61 [0.44, 0.79]	<.001	21.49 (0.57)	0.45 [0.31, 0.60]	<.001	0.08 [-0.14, 0.30]	.486

Generalized anxiety								
Baseline	8.87 (0.29)			8.98 (0.28)				
Follow up	6.13 (0.35)	0.70 [0.53, 0.86]	<.001	6.05 (0.33)	0.64 [0.50, 0.78]	<.001	-0.04 [-0.26, 0.17]	.684
Depression symptoms								
Baseline	8.60 (0.33)			9.21 (0.32)				
Follow up	5.60 (0.35)	0.72 [0.56, 0.88]	<.001	6.65 (0.33)	0.57 [0.43, 0.71]	<.001	0.10 [-0.11, 0.31]	.347
Mechanism Variables								
Distress Tolerance								
Baseline	21.38 (0.38)			21.29 (0.37)				
Postintervention	24.01 (0.32)	0.60 [0.47, 0.73]	<.001	23.24 (0.30)	0.43 [0.31, 0.54]	<.001	0.15 [-0.03, 0.32]	.095
Follow up	23.20 (0.36)	0.41 [0.25, 0.57]	<.001	22.53 (0.34)	0.27 [0.13, 0.42]	<.001	0.13 [-0.09, 0.34]	.250
Self-efficacy								
Baseline	5.90 (0.13)			5.83 (0.13)				
Postintervention	7.23 (0.13)	0.92 [0.76, 1.09]	<.001	7.02 (0.12)	0.68 [0.55, 0.80]	<.001	0.09 [-0.12, 0.29]	.397
Follow up	7.06 (0.13)	0.81 [0.62, 1.00]	<.001	6.67 (0.13)	0.48 [0.32, 0.63]	<.001	0.20 [-0.04, 0.45]	.096

Note. Within-group change p-values were adjusted using the Holm–Bonferroni correction within each variable to control for multiple comparisons.

Table S9. Fixed Effects Estimates from the Linear Mixed-Effects Model Based on the Complete-Case Sample in Study 2.

	B	SE	*p*
Primary Outcomes			
Academic anxiety			
Intercept	12.16	0.45	<.001
Group	0.46	0.29	.110
$Time_{baseline,\ follow}$	-2.82	0.29	<.001
Group × $Time_{baseline,\ follow}$	-1.00	0.42	.019
Negative emotion at baseline	0.12	0.02	<.001
Procrastination			
Intercept	15.68	0.72	<.001
Group	0.91	0.49	.063
$Time_{baseline,\ follow}$	-2.96	0.41	<.001
Group × $Time_{baseline,\ follow}$	-1.49	0.59	.012
Negative emotion at baseline	0.10	0.02	<.001
State anxiety			
Intercept	29.45	1.33	<.001
Group	0.00	0.93	1.000
$Time_{baseline,\ post}$	-9.12	0.83	<.001
$Time_{baseline,\ follow}$	-6.53	0.93	<.001
Group × $Time_{baseline,\ post}$	-2.30	1.20	.057
Group × $Time_{baseline,\ follow}$	-1.66	1.34	.217
Negative emotion at baseline	0.77	0.05	<.001
Secondary Outcomes			
Positive emotion			
Intercept	30.00	1.18	<.001
Group	0.27	0.77	.727
$Time_{baseline,\ post}$	2.27	0.51	<.001
$Time_{baseline,\ follow}$	2.13	0.52	<.001
Group × $Time_{baseline,\ post}$	0.73	0.74	.325
Group × $Time_{baseline,\ follow}$	0.16	0.76	.837
Negative emotion at baseline	-0.07	0.04	.105
Negative emotion			
Intercept	25.44	0.66	<.001
Group	-2.27	0.95	.018
$Time_{baseline,\ post}$	-5.99	0.56	<.001
$Time_{baseline,\ follow}$	-3.94	0.64	<.001
Group × $Time_{baseline,\ post}$	-0.78	0.81	.337
Group × $Time_{baseline,\ follow}$	-0.65	0.93	.486
Generalized anxiety			
Intercept	2.28	0.63	<.001
Group	-0.11	0.40	.794

$Time_{baseline,\ follow}$	-2.93	0.33	<.001
Group × $Time_{baseline,\ follow}$	0.19	0.47	.684
Negative emotion at baseline	0.27	0.02	<.001
Depressive symptoms			
Intercept	4.19	0.70	<.001
Group	-0.61	0.46	.185
$Time_{baseline,\ follow}$	-2.55	0.32	<.001
Group × $Time_{baseline,\ follow}$	-0.44	0.47	.347
Negative emotion at baseline	0.21	0.02	<.001
Mechanism Variables			
Distress tolerance			
Intercept	21.76	0.72	<.001
Group	0.09	0.53	.866
$Time_{baseline,\ post}$	1.95	0.28	<.001
$Time_{baseline,\ follow}$	1.25	0.34	<.001
Group × $Time_{baseline,\ post}$	0.67	0.40	.095
Group × $Time_{baseline,\ follow}$	0.57	0.50	.250
Negative emotion at baseline	-0.02	0.02	.429
Self-efficacy			
Intercept	6.73	0.26	<.001
Group	0.07	0.18	.718
$Time_{baseline,\ post}$	1.19	0.11	<.001
$Time_{baseline,\ follow}$	0.84	0.13	<.001
Group × $Time_{baseline,\ post}$	0.14	0.17	.397
Group × $Time_{baseline,\ follow}$	0.33	0.20	.096
Negative emotion at baseline	-0.04	0.01	<.001

Note. The Group variable was coded with the control group as the reference category; The Time variable was coded with baseline as the reference point.

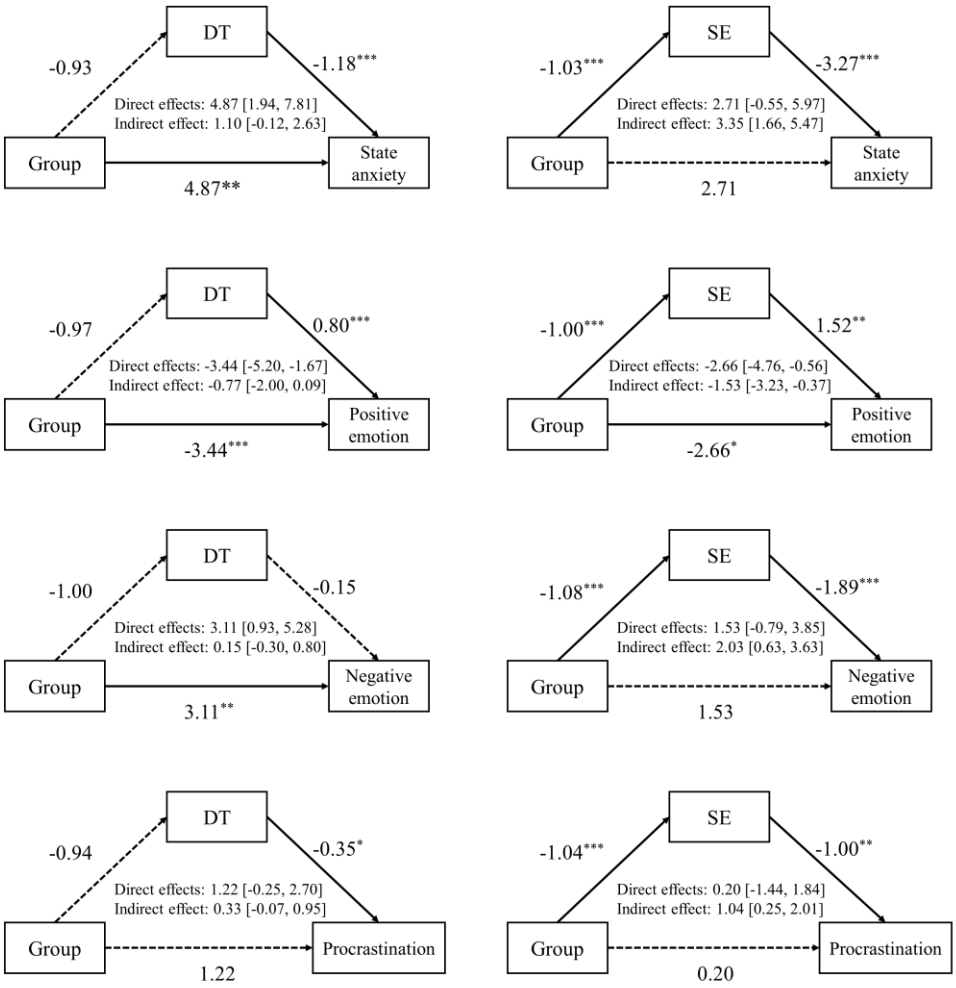

Figure S1. Mediation Models Examining the Indirect Effects of Distress Tolerance and Self-Efficacy on Study Outcomes in Study 1. Unstandardized path coefficients are presented. Solid lines indicate statistically significant paths ($p < .05$). Direct and indirect effects are reported with 95% confidence intervals (CIs). SE = Self-Efficacy; DT = Distress Tolerance. The models used postintervention levels of SE and DT as mediators. State anxiety, positive emotion, and negative emotion represent postintervention levels, whereas procrastination represents the follow-up level. Each model controlled for the corresponding baseline levels of the mediator and outcome variable, as well as paternal education and baseline depressive symptoms. Group was coded as 1 = intervention group and 2 = control group. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

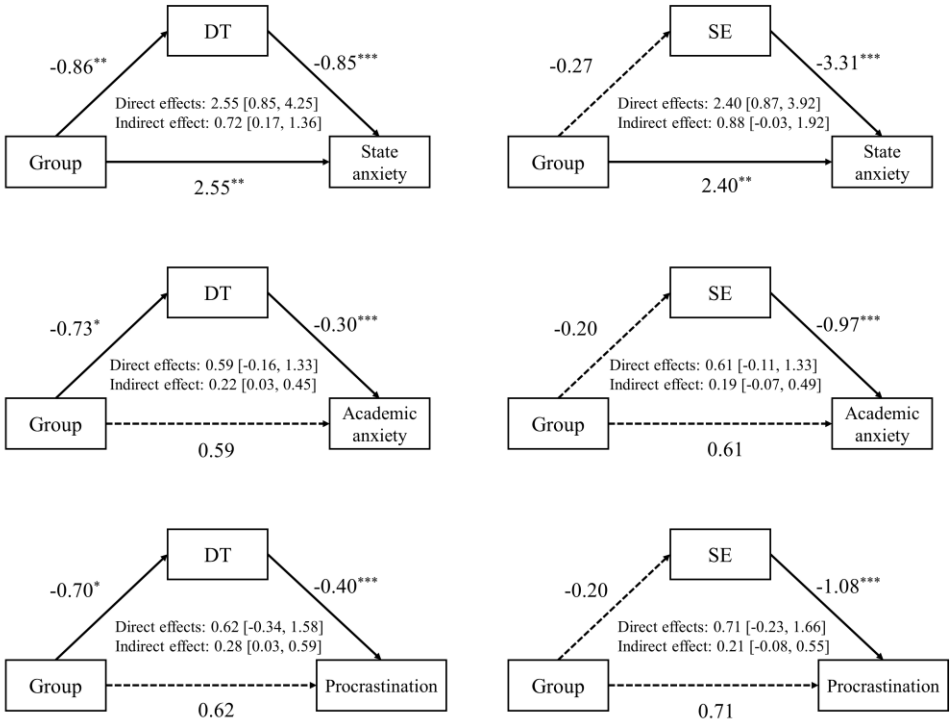

Figure S2. Mediation Models Examining the Indirect Effects of Distress Tolerance and Self-Efficacy on Study Outcomes in Study 2. Unstandardized path coefficients are presented. Solid lines indicate statistically significant paths ($p < .05$). Direct and indirect effects are reported with 95% confidence intervals (CIs). SE = Self-Efficacy; DT = Distress Tolerance. The models used postintervention levels of SE and DT as mediators. State anxiety represents postintervention level, whereas academic anxiety and procrastination represent the follow-up levels. Each model controlled for the corresponding baseline levels of the mediator and outcome variable, as well as baseline negative emotion. Group was coded as 1 = intervention group and 2 = control group. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Table S10. Internal consistency (Cronbach's α) of Study 2 measures that were also used in Study 1 at baseline, postintervention, and follow-up.

Measures	Cronbach's α		
	Baseline	Postintervention	Follow up
Psychotherapy Side Effects Scale (PSES)		0.70	
State–Trait Anxiety Inventory —State Anxiety Subscale (STAI-S)	0.94	0.92	0.94
Pure Procrastination Scale (PPS)	0.85		0.86
Positive and Negative Affect Schedule (PANAS)			
Positive Affect	0.87	0.89	0.89
Negative Affect	0.92	0.92	0.92
Patient Health Questionnaire-9 (PHQ-9)	0.80		0.84
Generalized Anxiety Disorder Scale-7 (GAD-7)	0.88		0.90
Multidimensional Experiential Avoidance Questionnaire–30 (MEAQ-30)			
Distress Endurance Subscale	0.89	0.88	0.90
Client Satisfaction Questionnaire-8 (CSQ-8)		0.88	

Validation results of the 5-item short form of the Academic Anxiety Scale (AAS) used in study 2

We adopted a combined data-driven and theory-based approach for item selection. Data were collected from an independent validation sample of 279 university students. We conducted an exploratory factor analysis (EFA) on the 11 items of the AAS using oblimin rotation based on the total validation sample of 279 participants. Although the original scale was developed as a unidimensional measure, we forced the extraction of two factors to identify potential item clusters that might reflect distinct aspects of academic anxiety.

The two-factor model derived from the exploratory factor analysis demonstrated excellent fit, CFI = 0.966, TLI = 0.944 and SRMR = 0.028. The detailed factor loadings are presented in Table S11. Factor 1 showed significant loadings for Items 1, 2, 3, 4, and 10, which we summarized as *performance-related academic anxiety*, including worries about meeting expectations, concerns about task completion, lower academic confidence, avoidance tendencies, and task-evoked physiological distress. Factor 2 showed significant loadings for Items 2 and 4 through 11, which we summarized as *contextual and social-evaluative school anxiety*, such as classroom fear, intimidation by instructors, concerns about peer evaluation, and broader school-related worries. Items 2, 4, and 10 significantly loaded on both factors, indicating that these experiences can be elicited by both performance-related academic demands and broader school or social-evaluative contexts.

For outcome measurement, we selected Items 1, 2, 3, 4, and 10, which loaded on Factor 1. These items align with the core mechanisms of our exposure intervention—reducing avoidance, modifying maladaptive threat expectations, and increasing tolerance of distress by directly confronting anxiety-eliciting academic tasks—and therefore provide the most theoretically relevant indicators of change for this intervention.

A confirmatory factor analysis (CFA) was subsequently conducted on these five items within the same validation sample (N = 279). The one-factor model demonstrated excellent fit CFI = 0.966, TLI = 0.932 and SRMR = 0.032. The short form also showed good internal consistency (Cronbach's $\alpha = 0.85$).

Table S11. Factor Loadings for the 11 Items of the Academic Anxiety Scale (AAS) from the Two-Factor Exploratory Factor Analysis

Items	Factor 1	Factor 2
1. I often worry that my best is not as good as expected in school.	0.785*	-0.049
2. I tend to put off doing school work because it stresses me.	0.444*	0.334*
3. I often worry that I am not doing assignments properly.	0.705*	0.054
4. I am less confident about school than my classmates.	0.397*	0.378*
5. I have a sense of dread when I am in my classrooms.	-0.046	0.854*
6. I tend to find my instructors intimidating.	-0.097	0.744*
7. I spend much of my time at school worrying about what is next.	-0.001	0.805*
8. There is something about school that scares me.	0.057	0.737*
9. I'm concerned about what my classmates think about my abilities.	0.183	0.539*
10. I often feel sick when I need to work on a major class assignment.	0.303*	0.562*
11. I have a hard time handling school responsibilities.	-0.018	0.751*

Note. * indicates that the loading was statistically significant at the $p < .05$ level.

CONSORT 2025 checklist of information to include when reporting a randomised trial*

Section / Topic	No	CONSORT 2025 checklist item description	Reported on page no.
Title and abstract			
Title and structured abstract	1a	Identification as a randomised trial	3
	1b	Structured summary of the trial design, methods, results, and conclusions	3
Open science			
Trial registration	2	Name of trial registry, identifying number (with URL) and date of registration	27 & 41
Protocol and statistical analysis plan	3	Where the trial protocol and statistical analysis plan can be accessed	27 & 41
Data sharing	4	Where and how the individual de-identified participant data (including data dictionary), statistical code and any other materials can be accessed	46
Funding and conflicts of interest	5a	Sources of funding and other support (e.g., supply of drugs), and role of funders in the design, conduct, analysis and reporting of the trial	55
	5b	Financial and other conflicts of interest of the manuscript authors	55
Introduction			
Background and rationale	6	Scientific background and rationale	4-7
Objectives	7	Specific objectives related to benefits and harms	6-7
Methods			
Patient and public involvement	8	Details of patient or public involvement in the design, conduct and reporting of the trial	N/A
Trial design	9	Description of trial design including type of trial (e.g., parallel group, crossover), allocation ratio, and framework (e.g., superiority, equivalence, non-inferiority, exploratory)	27-28 & 41-42
Changes to trial protocol	10	Important changes to the trial after it commenced including any outcomes or analyses that were not prespecified, with reason	34
Trial setting	11	Settings (e.g., community, hospital) and locations (e.g., countries, sites) where the trial was conducted	27 & 41
Eligibility criteria	12a	Eligibility criteria for participants	27-28 & 41-42
	12b	If applicable, eligibility criteria for sites and for individuals delivering the interventions (e.g., surgeons, physiotherapists)	30

Intervention and comparator	13	Intervention and comparator with sufficient details to allow replication. If relevant, where additional materials describing the intervention and comparator (e.g., intervention manual) can be accessed	28-33 & 42-43
Outcomes	14	Pre-specified primary and secondary outcomes, including the specific measurement variable (e.g., systolic blood pressure), analysis metric (e.g., change from baseline, final value, time to event), method of aggregation (e.g., median, proportion), and time point for each outcome	35-38 & 44-45
Harms	15	How harms were defined and assessed (e.g., systematically, non-systematically)	35
Sample size	16a	How sample size was determined, including all assumptions supporting the sample size calculation	28 & 42
	16b	Explanation of any interim analyses and stopping guidelines	N/A
Randomisation:			
Sequence generation	17a	Who generated the random allocation sequence and the method used	28 & 42
	17b	Type of randomisation and details of any restriction (e.g., stratification, blocking and block size)	28 & 42
Allocation concealment mechanism	18	Mechanism used to implement the random allocation sequence (e.g., central computer/telephone; sequentially numbered, opaque, sealed containers), describing any steps to conceal the sequence until interventions were assigned	28
Implementation	19	Whether the personnel who enrolled and those who assigned participants to the interventions had access to the random allocation sequence	28
Blinding	20a	Who was blinded after assignment to interventions (e.g., participants, care providers, outcome assessors, data analysts)	27
	20b	If blinded, how blinding was achieved and description of the similarity of interventions	N/A (unblinded)
Statistical methods	21a	Statistical methods used to compare groups for primary and secondary outcomes, including harms	38-40 &45
	21b	Definition of who is included in each analysis (e.g., all randomised participants), and in which group	38
	21c	How missing data were handled in the analysis	38
	21d	Methods for any additional analyses (e.g., subgroup and sensitivity analyses), distinguishing prespecified from post-hoc	39 & 40
Results			
Participant flow, including flow diagram	22a	For each group, the numbers of participants who were randomly assigned, received intended intervention, and were analysed for the primary outcome	Fig. 1 & Fig. 2
	22b	For each group, losses and exclusions after randomisation, together with reasons	Fig. 1 & Fig. 2
Recruitment	23a	Dates defining the periods of recruitment and follow-up for outcomes of benefits and harms	27, 34, 41 & 43
	23b	If relevant, why the trial ended or was stopped	N/A

Intervention and comparator delivery	24a	Intervention and comparator as they were actually administered (e.g., where appropriate, who delivered the intervention/comparator, how participants adhered, whether they were delivered as intended [fidelity])	33,42-43, Fig. 1 & Fig. 2
	24b	Concomitant care received during the trial for each group	27-28
Baseline data	25	A table showing baseline demographic and clinical characteristics for each group	Table 1
Numbers analysed, outcomes and estimation	26	For each primary and secondary outcome, by group: ● the number of participants included in the analysis ● the number of participants with available data at the outcome time point ● result for each group, and the estimated effect size and its precision (such as 95% confidence interval) ● for binary outcomes, presentation of both absolute and relative effect size	8-17 Table 2 & 3
Harms	27	All harms or unintended events in each group	8-9 & 13-14
Ancillary analyses	28	Any other analyses performed, including subgroup and sensitivity analyses, distinguishing pre-specified from post-hoc	9, 11-12 & 16-17
Discussion			
Interpretation	29	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	18-26
Limitations	30	Trial limitations, addressing sources of potential bias, imprecision, generalisability, and, if relevant, multiplicity of analyses	24-25

Citation: Hopewell S, Chan AW, Collins GS, Hróbjartsson A, Moher D, Schulz KF, et al. CONSORT 2025 Statement: updated guideline for reporting randomised trials. BMJ. 2025; 388:e081123. https://dx.doi.org/10.1136/bmj-2024-081123.