

Supplementary Table 1 Domain composition and scoring structure of the psychological well-being instrument

Domain	Component items	Score used in analyses
Mental fatigue	Item1, Item2	Domain score
Workload	Item3, Item7	Domain score
Stress level	Item5	Domain score
Self-efficacy	Item4, Item8	Domain score
Emotional state	Item6_1, Item6_2, Item6_3	Domain score
Overall psychological well-being	Aggregated from the five domain scores	Overall score

Notes. Scores were generated according to the predefined operational scoring protocol used in the monitoring system.

Supplementary Table 2 Internal consistency estimates for the multi-item domains of the psychological well-being instrument

Domain	Number of items	Cronbach's α
Mental fatigue	2	0.848
Workload	2	0.868
Self-efficacy	2	0.833
Emotional state	3	0.783

Notes. Internal consistency was summarized for the multi-item domains. The overall psychological well-being score was used as a composite operational index in the predictive analyses.

Supplementary Table 3 Task-specific Poisson-model results for the overall psychological well-being score, with false-discovery-rate adjustment

Task	Estimate	95% CI	<i>p</i>	FDR-q
Balloon	0.385	0.193 to 0.771	0.007	0.047
BallSport	0.632	0.364 to 1.072	0.097	0.226
Stroop accuracy	0.729	0.498 to 1.088	0.115	0.226
Ligature test	0.690	0.423 to 1.104	0.131	0.226
Stroop time	0.750	0.508 to 1.134	0.162	0.226
Digit memory	0.929	0.567 to 1.477	0.764	0.892
Target tracking	1.013	0.653 to 1.574	0.955	0.955

Notes. Estimates < 1 indicate a protective association of higher psychological well-being with lower risk of non-full-score events.

Supplementary Table 4 Record-level generalized linear mixed models as robustness analyses

Task	OR for overall well-being	95% CI	<i>p</i>	Note
BallSport	0.755	0.150 to 3.800	0.733	Supplementary robustness model
Target tracking	1.469	0.439 to 4.922	0.533	Supplementary robustness model
Digit memory	0.592	0.113 to 3.092	0.534	Supplementary robustness model
Balloon	Not reliably estimated	Not reliably estimated	—	Estimate unstable because of sparse events; interpret cautiously
Ligature test	0.658	0.391 to 1.105	0.114	Supplementary robustness model
Stroop time	0.335	0.086 to 1.307	0.115	Supplementary robustness model
Stroop accuracy	0.320	0.083 to 1.242	0.100	Supplementary robustness model

Notes. These mixed models were fitted as supplementary robustness analyses. Some task-specific estimates should be interpreted cautiously because the number of non-full-score events was limited in several subsamples. The Balloon-task estimate was not reported quantitatively because sparse events led to unstable model estimation.