Supplementary Online Materials (SOM)

Attitudinal Effects of Stimulus Co-occurrence and Stimulus Relations:

Sleep Supports Propositional Learning via Memory Consolidation

1. Chronotype and Time 1 alertness did not differ between conditions

Participants' chronotypes were not significantly different across sleep and wake conditions, χ^2 (2) = 0.335, $p = .846$ (Table S1).

To address the concern that chronotype may have contributed to the sleep effect, we conducted exploratory analyses using condition (sleep vs. wake) and chronotype as between-participant factors. Following Bodenhausen (1990), we used a median split of rMEQ scores to categorize participants into morning vs. evening types. If chronotype significantly influenced memory preservation from T1 to T2, then we should expect a significant chronotype-by-condition interaction, as in Bodenhausen (1990).

However, the chronotype-by-condition interaction was not significant for any of the five memory indices: $F(1, 196) = 0.12$, $p = .727$ for US_identity, $F(1, 196) < 0.01$, $p = .999$ for US_Valence, $F(1, 196) = 0.36$, $p = .547$ for Relation, $F(1, 196) = 0.66$, $p = .419$ for US_identity+relation, and $F(1, 196) = 0.09$, $p = .767$ for US_valence+relation.

Importantly the critical Sleep vs. Wake condition effects remained significant for US_identity, $F(1, 196) = 4.91$, $p = .028$, and US_identity+relation, $F(1, 196) = 4.37$, $p = .038$. These analyses suggest that chronotype did not play a significant role in the observed effect.

To control for time-of-day effects, we also assessed participants' alertness levels with the Stanford Sleepiness Scale (SSS). There was no significant difference in alertness between the sleep and the wake condition at Time 1, $t(198) = 0.05$, $p = .958$, $d = 0.01$, BF_{10}=0.15, with BF indicating that the data is 6 times more likely under the null model than to the alternative model. However, at Time 2, wake participants did report a significantly lower level of alertness at Time 2, $t(198) = 4.56$, $p < .001$, $d = 0.65$. To statistically control for the influence of Time 2 alertness, we added this variable as a covariate in an ANCOVA model, using sleep vs. wake conditions to predict memory preservation scores for US_identity and US_identity+relation. The ANCOVA results were consistent with what was reported in the main text: participants in the sleep condition still showed significantly higher preservation scores on US_identity, $F(1, 197) = 4.10$, $p = .040$, $\eta^2_p = .020$, and on US_identity+relation, $F(1, 197) = 3.97$, $p = .048$, η^2_p= .020. Thus, the between-condition differences in T2 memory remained significant after controlling for T2 alertness levels.

Table S1
Participants' chronotype and self-reported alertness (mean ± S.D., alertness was assessed at the end of each session).

	Sleep (*n* =100)	Wake (*n*= 100)
Chronotype (rMEQ)		
Morning Type	12	11
Evening Type	30	27
Neutral Type	58	62
Alertness (SSS)		
Time 1	3.22 (1.23)	3.29 (1.34)
Time 2	2.45 (1.16)	3.28 (1.32)

Note. rMEQ = Reduced Morningness-Eveningness Questionnaire; SSS = Stanford Sleepiness Scale.

2. Memory results were robust when Time 1 memory differences were controlled

At Time 1, there were no significant between-condition differences in all five memory scores. It is worth mentioning that for memory for US identity, participants in the wake condition tended to recall more correct pairs than participants in the sleep condition, an effect approaching significance, $t(198) = 1.89$, $p = .060$, $d = 0.27$. To control for Time 1 differences in memory, we conducted a between-subject (wake vs. sleep) ANCOVA using Time 1 memory as a covariate on Time 2 memory scores. The results confirmed the results of analyses reported in the main text, such that at Time 2, sleep led to better memory for US identity than wake, $F(1, 197) = 8.99$, $p = .003$, $\eta^2_p = .044$.

3. Memory results were robust when potential outliers were removed

A visual inspection of Figure 3 suggests that both conditions included potential outliers. Here, we re-ran all analyses excluding outliers identified via the interquartile range (IQR) for each memory score. All results on memory preservation were robust against outlier removal: participants in the sleep condition showed significantly higher preservation scores for US_identity, $t(180) = 3.41$, $p < .001$, $d = 0.51$, and US_identity+relation, $t(182) = 3.35$, $p < .001$, $d = 0.49$. In addition, participants in the sleep condition performed better than participants in the wake condition on US_valence after removing potential outliers, $t(182) = 2.19$, $p = .030$, $d = 0.32$.

4. RCB Modeling results were robust when using different RT cutoffs in the evaluative choice task

The RCB model's *R* parameter quantifies the impact of CS-US relations on evaluative choices. A previous study found that *R* were greater when participants had more time during the evaluative choice task (Heycke & Gawronski, 2020, Experiment 4). We therefore tested whether our RCB modeling results are robust when using different RT cutoffs.

At the trial-level, we imposed three reaction time criteria (RT > 200, 300, and 500msec). We excluded 182 trials (0.83% of total trials) with RTs < 200msec, 811 trials (3.72% of total trials) with RTs < 300msec, and 4089 trials (18.74% of total trials) with RTs < 500msec.

We next repeated the RCB modeling analysis using each RT cutoff criterion. For all three cutoff criteria, the baseline models fit the data well (RT > 200msec: $G^2(4) = 4.46$, $p = .348$; RT > 300msec: $G^2(4) = 5.10$, $p = .277$; RT > 500msec: $G^2(4) = 6.57$, $p = .160$, see Table S2 for details of parameter estimates and their 95% CIs).

Consistent with the results reported in the main text, we found that the *R* parameter significantly increased from Time 1 to Time 2 in the sleep condition regardless of the employed cutoff criteria (RT > 200msec: $\Delta G^2(1) = 4.90$, $p = .027$; RT > 300msec: $\Delta G^2(1) = 4.82$, $p = .028$; RT > 500msec: $\Delta G^2(1) = 4.68$, $p = .030$), but remained the same in the wake condition (RT > 200msec: $\Delta G^2(1) = 1.16$, $p = .281$; RT > 300msec: $\Delta G^2(1) = 0.743$, $p = .389$; RT > 500msec: $\Delta G^2(1) = 1.49$, $p = .222$).

Consistent with the results reported in the main text, we did not find significant Time 1 vs. Time 2 differences on the *C* parameter in both the sleep condition (RT > 200msec: $\Delta G^2(1) = 0.11$, $p = .739$; RT > 300msec: $\Delta G^2(1) = 0.18$, $p = .668$; RT > 500msec: $\Delta G^2(1) = 1.16$, $p = .281$), and the wake condition (RT > 200msec: $\Delta G^2(1) = 0.09$, $p = .763$; RT > 300msec: $\Delta G^2(1) = 0.12$, $p = .726$; RT > 500msec: $\Delta G^2(1) = 0.03$, $p = .865$).

Consistent with the results reported in the main text, we found a significant effect of Time on the *B* parameter in both the sleep condition (RT > 200msec: $\Delta G^2(1) = 14.01$, $p < .001$; RT > 300msec: $\Delta G^2(1) = 16.44$, $p < .001$; RT > 500msec: $\Delta G^2(1) = 14.20$, $p < .001$), and the wake condition (RT > 200msec: $\Delta G^2(1) = 14.27$, $p < .001$; RT > 300msec: $\Delta G^2(1) = 17.94$, $p < .001$; RT > 500msec: $\Delta G^2(1) = 14.84$, $p < .001$).

Descriptively, both the *R* and the *C* parameter became larger after excluding shorter RTs. The *B* parameters decreased as we excluded shorter RTs: when participants took more time to make a judgment, they showed less general response biases.

Table S2
Parameter Estimates of the RCB Model as a Function of Time, Experimental Condition and Trial RT.

	Sleep				Wake			
	Time 1		Time 2		Time 1		Time 2	
	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
R								
All	.19	[.16, .21]	.22	[.20, .25]	.17	[.14, .20]	.19	[.16, .22]
RT > 200ms	.19	[.16, .21]	.23	[.20, .25]	.17	[.14, .20]	.19	[.16, .22]
RT > 300ms	.19	[.17, .22]	.24	[.21, .26]	.18	[.15, .21]	.20	[.17, .22]
RT > 500ms	.23	[.20, .25]	.27	[.24, .30]	.19	[.16, .22]	.22	[.19, .25]
C								
All	.14	[.10, .17]	.14	[.11, .18]	.11	[.08, .14]	.10	[.07, .13]
RT > 200ms	.14	[.10, .17]	.15	[.11, .18]	.11	[.08, .14]	.10	[.07, .14]
RT > 300ms	.14	[.11, .17]	.15	[.12, .19]	.11	[.08, .15]	.10	[.07, .14]

RT > 500ms	.16	[.12, .19]	.18	[.14, .22]	.13	[.10, .17]	.13	[.09, .16]
B								
All	.59	[.57, .61]	.54	[.52, .56]	.55	[.53, .57]	.50	[.49, .52]
RT > 200ms	.59	[.57, .61]	.54	[.52, .56]	.55	[.54, .57]	.50	[.49, .52]
RT > 300ms	.59	[.57, .61]	.53	[.51, .55]	.56	[.54, .58]	.50	[.49, .52]
RT > 500ms	.58	[.56, .60]	.52	[.49, .54]	.55	[.54, .58]	.49	[.47, .52]