**Experiment 1**

**Results**

Solution rates were analysed by means of generalised linear mixed effects (glme) models, using the “glmer” function from the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) and the binomial family, within the R environment for statistical computing (R Core Team, 2015). CRAT problem difficulty and sound conditions were coded as fixed factors and specified using the function “contr.sdif” from the MASS package (Venables & Ripley, 2002). Our general approach was to include the full random effects structure, with both random intercept and slopes for all the fixed effects across subjects and items (Barr, Levy, Scheepers, & Tily, 2013) and specify 10000 iterations with the optimizer “bobyqa” for glme models. If a model failed to converge and overfitted the data, we systematically trimmed the model starting with items and then participants, by first removing correlations between factors and then interactions. If a model still failed to converge and overfitted the data, we removed each random slope, and as last resort, we removed one random structure. The p values were estimated using the lmerTest package (Kuznetsova, Brockhoff, Christensen, 2017). Fixed effect estimations are summarised in Table X.

Self-reported solution strategies scores on the four-point scale were standardised so that “1” always indicated an analytic solution, and “4” always indicated an insight solution. These strategies were treated as ordinal responses and analysed with cumulative link mixed (clm) modelling in R, using the “clmm” function from the ordinal package (Christensen, 2015). Difficulty and sound conditions were specified as fixed factors using the default contr.treatment and using as baseline the Steady-State Sound condition for Experiment 1, and the Meaningless Sound condition for Experiment 3. We used this sound conditions as baseline to allow us to perform theoretically meaningful contrasts, and have results easily comparable with the analyses on the solution rates. The models initially included a full random effects structure (Barr et al., 2013), with a random intercept and random slopes per subject and item. If a model failed to converge, we increased the number of inner optimizations using the default optimizer “ucminf”, and then systematically trim the random structure by first removing correlations between factors and then interactions. The significance of all effects was assessed using the “summary” command, with the default p-values based on the Wald statistic. If any interaction reached significance, post hoc comparisons of estimated marginal means with Tukey adjustment were performed with emmeans package (Lenth et al., 2018) in R. The model results are shown in Table XI.

Table X

*Generalised Linear Mixed Effects Models for All Experiments*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | Experiment 1 | | | | |
|  |  | Solution Rate | | | |  |
|  |  | *b* | *SE* | *z* | *Sign*. | |
| Factor |  |  |  |  |  |  |
| Intercept |  | -0.538 | 0.116 | -4.636 | 0.000 | \*\*\* |
| Difficult - Easy |  | -0.637 | 0.113 | -5.651 | 0.000 | \*\*\* |
| Steady-State - Quiet |  | -0.089 | 0.101 | -0.880 | 0.379 |  |
| Changing-State - Steady-State |  | -0.535 | 0.105 | -5.082 | 0.000 | \*\*\* |
| (Difficult – Easy) ´ (Steady-State – Quiet) |  | -0.290 | 0.202 | -1.434 | 0.152 |  |
| (Difficult – Easy) ´ (Changing-State - Steady-State) |  | 0.050 | 0.210 | 0.240 | 0.810 |  |
|  |  | Experiment 2 | | | | |
|  |  | Solution Rate | | | |  |
|  |  | *b* | *SE* | *z* | *Sign*. | |
| Factor |  |  |  |  |  |  |
| Intercept |  | -0.752 | 0.115 | -6.538 | 0.000 | \*\*\* |
| Difficult - Easy |  | -0.926 | 0.102 | -9.115 | 0.000 | \*\*\* |
| Steady-State - Quiet |  | -0.050 | 0.120 | -0.415 | 0.678 |  |
| Changing-State - Steady-State |  | -0.528 | 0.125 | -4.213 | 0.000 | \*\*\* |
| (Difficult – Easy) ´ (Steady-State – Quiet) |  | 0.265 | 0.240 | 1.108 | 0.268 |  |
| (Difficult – Easy) ´ (Changing-State - Steady-State) |  | -0.046 | 0.250 | -0.184 | 0.854 |  |
|  |  | Experiment 3 | | | | |
|  |  | Solution Rate | | | |  |
|  |  | *b* | *SE* | *z* | *Sign*. | |
| Factor |  |  |  |  |  |  |
| Intercept |  | -0.811 | 0.149 | -5.435 | 0.000 | \*\*\* |
| Difficult - Easy |  | -1.228 | 0.190 | -6.454 | 0.000 | \*\*\* |
| Meaningless - Quiet |  | -0.410 | 0.131 | -3.125 | 0.002 | \*\* |
| Meaningful - Meaningless |  | -0.337 | 0.140 | -2.402 | 0.016 | \* |
| (Difficult – Easy) ´ (Meaningless - Quiet) |  | -0.494 | 0.262 | -1.881 | 0.060 | . |
| (Difficult – Easy) ´ (Meaningful - Meaningless) |  | -0.011 | 0.281 | -0.040 | 0.968 |  |

Table XI

*Cumulative Link Mixed Effects Models for Experiments 1 and 3*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | Experiment 1 | | | | |
|  |  | Self-Reported Solution Strategies | | | |  |
|  |  | *b* | *SE* | *z* | *Sign*. | |
| Factor |  |  |  |  |  |  |
| Difficult |  | 0.144 | 0.301 | 0.478 | 0.632 |  |
|  |  |  |  |  |  |  |
| Quiet |  | 0.514 | 0.295 | 1.739 | 0.082 | . |
|  |  |  |  |  |  |  |
| Changing-State |  | -0.003 | 0.407 | -0.007 | 0.994 |  |
|  |  |  |  |  |  |  |
| Difficult:Quiet |  | -0.538 | 0.405 | -1.329 | 0.184 |  |
|  |  |  |  |  |  |  |
| Difficult:Changing-State |  | -0.415 | 0.447 | -0.929 | 0.353 |  |
|  |  | Experiment 3 | | | | |
|  |  | Self-Reported Solution Strategies | | | |  |
|  |  | *b* | *SE* | *z* | *Sign*. | |
| Factor |  |  |  |  |  |  |
| Difficult |  | -0.840 | 0.289 | -2.903 | 0.004 | \*\* |
|  |  |  |  |  |  |  |
| Quiet |  | 0.045 | 0.223 | 0.203 | 0.839 |  |
|  |  |  |  |  |  |  |
| Meaningful |  | -0.142 | 0.243 | -0.582 | 0.560 |  |
|  |  |  |  |  |  |  |
| Difficult:Quiet |  | 0.081 | 0.383 | 0.212 | 0.832 |  |
|  |  |  |  |  |  |  |
| Difficult:Meaningful |  | 1.242 | 0.449 | 2.768 | 0.006 | \*\* |

***Solution Rates***

Figure 1 displays the mean solution rates according to each level of Sound and each level of CRAT problem difficulty. As can be seen, the Changing-state speech condition resulted in fewer correct responses in comparison to either the Steady-state speech or Quiet conditions. In addition, solution rates for Easy CRAT problems were higher than for Difficult CRAT problems, but problem difficulty did not appear to moderate the effect of Sound on CRAT performance.

The final generalized linear mixed effects model for Experiment 1 reached convergence with random intercepts for each subject and item. The analysis revealed a significant main effect of CRAT problem difficulty. As expected, solution rates were significantly higher for the Easy problems (*M* = .45, *SE* = .03) in comparison to the Difficult problems (*M* = .32, *SE* = .03). In addition, significantly more problems were solved in the Steady-state speech condition (*M* = .42, *SE* = .03) in comparison to the Changing-state speech condition (*M* = .31, *SE* = .03). There was no significant difference between the number of problems solved in the Quiet (*M* = .44, *SE* = .03) and the Steady-state speech conditions (*M* = .42, *SE* = .03). There was no significant CRAT problem difficulty interaction x Sound.

***Self-Reported Solution Strategies***

The final model for Experiment 1 included a random intercept for item and a random structure with both fixed effects and their correlation, but not their interaction, for subject.

The self-reported scores for solution strategy were collated for CRAT problems that had been solved correctly. The analysis indicated that the self-reported solution strategy was not impacted by any of the experimental conditions. There was no significant main effect of CRAT problem difficulty and Sound conditions did not differ from each other. In addition, the CRAT problem difficulty x Sound interaction was not significant.

**Experiment 2**

**Results**

As in Experiment 1, we analysed solution rates with glme models in R, using the “glmer” function from the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) and the binomial family, and the self-reported solution strategies with clm models in R with the “clmm” function from the ordinal package (Christensen, 2015). The same fixed effects and approach to random structure were used when modelling. The results of the models are summarised in Table X and Table XI.

***Solution Rates***

Figure 2 shows the mean solution rates according to each level of Sound and each level of CRAT problem difficulty. The data indicate a reduction in CRAT solution performance for the Changing-state tone condition in comparison to both the Steady-state tone condition and the Quiet condition.

The model that reached convergence included a random intercept for subjects only. The analysis revealed a significant main effect of CRAT problem difficulty, with significantly higher solution rates for the Easy CRAT problems (M = .44, SE = .03) in comparison to the Difficult CRAT problems (M = .24, SE = .03). Furthermore, significantly more problems were solved in the Steady-state tone condition (*M* = .36, *SE* = .04) in comparison to the Changing-state tone condition (*M* = .29, *SE* = .03). There was no significant difference in the solution rate for problems in the Quiet condition and Steady-state tone condition. There was no significant CRAT problem difficulty x Sound interaction.

**Experiment 3**

**Results**

Consistent with Experiment 1 and in Experiment 2, we analysed the data in R with glme models (“glmer” function from the lme4 package; Bates, Mächler, Bolker, & Walker, 2015) for the solution rates, and clm models ( “clmm” function from the ordinal package, Christensen, 2015) for the self-reported solution strategies. The same Difficulty and Sound conditions were treated as fixed effects, and the same general approach adopted to build and then trim the random structure. As in Experiment 1, post hoc comparisons of estimated marginal means with Tukey adjustment were performed with emmeans package (Lenth et al., 2018) in R on any significant interaction observed in the clm model. The results of the models are shown in Table X and Table XI.

***Solution Rates***

Figure 3 shows the pattern of findings for the solution rates in Experiment 3. It is apparent that meaningful speech appeared to be detrimental to CRAT performance relative to the meaningless speech and quiet conditions. In addition, the data seem to indicate that meaningless speech also led to worse CRAT performance than the quiet condition.

The final glme model that reached convergence without overfitting the data included a random intercept for both subject and item, and a random slope (CRAT problem difficulty) for subject. Consistent with Experiments 1 and 2, there was a significant main effect of CRAT problem difficulty, with significantly higher solution rates for the Easy CRAT problems (M = .46, SE = .03) in comparison to the Difficult CRAT problems (M = .22, SE = .03). In addition, significantly more CRAT problems were solved in the Quiet condition (*M* = .41, *SE* = .04) in comparison to the Meaningless speech condition (*M* = .34, *SE* = .03). There was also a significant difference between the type of speech, with significantly fewer problems solved in the Meaningful speech condition (*M* = .28, *SE* = .03) in comparison to the Meaningless speech condition (*M* = .34, *SE* = .03). Finally, the analysis revealed a marginal CRAT problem difficulty x Sound interaction, such that the effect of CRAT problem difficulty was stronger in the Meaningless speech condition compared to the Quiet condition.

***Self-Reported Solution Strategies***

The self-reported solution strategy scores were collated for CRAT problems that had been solved correctly. The final model for Experiment 3 included a random intercept for both participant and item, and one slope (Sound) for the participant random effects structure only. Contrary to our prediction solution strategies did not differ significantly depending on the Sound conditions. However, CRAT problem difficulty did give rise to a significant main effect on self-reported solution strategy scores, with Easy CRAT problems (M = 2.98, SE = .08) resulting in significantly higher insight responses compared to with Difficult CRAT problems (M = 2.80, SE = .09). There was also a reliable interaction between the CRAT problem difficulty and the Meaningless and Meaningful Sound conditions. Pairwise comparisons revealed that for the Easy CRAT problems, Sound did not have any effect on self-reported solution strategy responses, as solution rates did not differ significantly between the Meaningless speech and Quiet conditions (*b* = -0.045, *SE =*0.223, *z* = -0.203), and between the Meaningless and Meaningful speech (*b* = 0.142, *SE =*0.243, *z* = 0.582 ) conditions. However, for the Hard CRAT problems, Sound did have an effect on self-reported solution strategy responses. The self-reported rating scores significantly differed between the Meaningless and Meaningful speech (*b* = -0.9733, *SE* = 0.358, *z* = -2.717) conditions, resulting in significant XXXX . We did not observe any significant difference in the Difficult CRAT problems between the Meaningless speech and Quiet conditions (*b* = -0.1266, *SE* = 0.323, *z* = -0.392).