**SUPPLEMENTARY MATERIALS FOR THE EMBODIED GOD**

**1. Frequentist Analyses for Exp. 1 used for the Exp. 2 Pre-registration**

Frequentist paired-sample t-tests (two-tailed) revealed that for psychology items, accuracy was significantly higher for consistent (M = 97.8%, SD = 4.7%) than inconsistent (M = 92.0%, SD = 7.5%) items; t (126) = 8.40, p < 0.001, d = 0.74, 95% CI = [0.55, 0.94]. Similarly, for physicality items, accuracy was significantly higher for consistent (M = 94.4, SD = 11.5%) than inconsistent (M = 89.6%, SD = 14.5%) items; t (126) = 4.85, p < .001, d = 0.43, 95% CI = [0.25, 0.61].

**2. Frequentist analyses for Exp. 1, 3, and 4**

Here we report Frequentist analyses for Exp. 1, 3, and 4 to supplement the Bayesian analyses reported in the main text. As with the Bayesian analysis reported in the main text, data points that were above or below 3SD from the mean RT of each statement were removed, and RT interference scores were calculated using both correct and incorrect responses.

**2.1. Exp. 1**

We used a Frequentist paired-samples t-test to evaluate whether the accuracy and RT interferences scores on psychology and physicality items were the same or different. For accuracy there was no significant difference between physicality (*M* = 4.74, *SD* = 11.01) and psychology (*M* = 5.76, *SD* = 7.73) interference scores; *t* (126) < 1.0, *p* > .25. For RT, interference scores were significantly higher for physicality (*M* = 269, *SD* = 513) than psychology (*M* = 116, *SD* = 330); *t* (126) = 2.94, *p* = .004, *d* = 0.26, 95% CI = [0.08, 0.44]. The RT interference score for physicality was more than twice as large as that for psychology.

We then used Frequentist one-sample t-tests to evaluate whether the accuracy and RT interferences scores on psychology and physicality were different from zero. Accuracy interference scores were significantly greater than zero for physicality, *t* (126) = 4.85, *p* < .001, *d* = 0.43, 95% CI = [0.25, 0.61], and psychology, *t* (126) = 8.40, *p* < .001, *d* = 0.75, 95% CI = [0.55, 0.94]. Similarly, RT interference scores were significantly greater than zero for physicality, *t* (126) = 5.91, *p* < .001, *d* = 0.52, 95% CI = [0.34, 0.71], and psychology, *t* (126) = 3.95, *p* < .001, *d* = 0.35, 95% CI = [0.17, 0.53].

**2.2. Exp. 3**

Response interference scores for accuracy and RT were each submitted to a 3 (Condition: God versus Holy Spirit versus Jesus) x 2 (Characteristic: Psychology versus Physicality) Frequentist mixed ANOVA to evaluate whether there were differences in response interference among the three members of the Trinity. For accuracy, results revealed a significant main effect of Characteristic [*F* (1, 216) = 4.82, *p* = .029, η²p = 0.02], but no main effect of Condition [*F* (2, 216) < 1.0, *p* > .250] nor an interaction [*F* (2, 216) = 2.68, *p* = .071]. A post-hoc test found that accuracy interference scores were greater for psychology (*M* = 7.73, *SD* = 9.91) than physicality (*M* = 5.54, *SD* = 12.73); *p* = .028, *d* = 0.15, 95% CI = [0.01, 0.28]. For RT, results revealed a significant main effect of Characteristic [*F* (1, 216) = 6.79, *p* = .010, η²p = 0.03], but no main effect of Condition [*F* (2, 216) < 1.0, *p* > .250] nor an interaction [*F* (2, 216) = 1.49, *p* = .227]. A post-hoc test found that RT interference scores were greater for physicality (*M* = 220, *SD* = 473) than psychology (*M* = 124, *SD* = 335); *p* = .011, *d* = 0.17, 95% CI = [0.04, 0.31].

We then used one-sample t-tests to evaluate whether the accuracy and RT interferences scores on psychology and physicality were different from zero for each of the three members of the Trinity. Accuracy and RT interferences scores for all three beings were greater than zero for physicality (**God**: accuracy *t* (67) = 4.10, *p* < .001, *d* = 0.50, 95% CI = [0.24, 0.75], RT *t* (67) = 5.51, *p* < .001, *d* = 0.67, 95% CI = [0.41, 0.93]; **Holy Spirit**: accuracy *t* (71) = 2.72, *p* = .008, *d* = 0.32, 95% CI = [0.08, 0.56], RT *t* (71) = 3.31, *p* = .001, *d* = 0.39, 95% CI = [0.15, 0.63]; **Jesus**: accuracy *t* (78) = 4.77, *p* < .001, *d* = 0.54, 95% CI = [0.30, 0.77], RT *t* (78) = 3.37, *p* = .001, *d* = 0.38, 95% CI = [0.15, 0.61]) and psychology (**God**: accuracy *t* (67) = 5.47, *p* < .001, *d* = 0.66, 95% CI = [0.40, 0.92], RT *t* (67) = 2.94, *p* = .005, *d* = 0.36, 95% CI = [0.11, 0.60]; **Holy Spirit**: accuracy *t* (71) = 8.15, *p* < .001, *d* = 0.96, 95% CI = [0.68, 1.24], RT *t* (71) = 3.73, *p* < .001, *d* = 0.44, 95% CI = [0.20, 0.63]; **Jesus**: accuracy *t* (78) = 6.41, *p* < .001, *d* = 0.72, 95% CI = [0.47, 0.97], RT *t* (78) = 2.80, *p* = .006, *d* = 0.32, 95% CI = [0.08, 0.54]).

**2.3. Exp. 4**

We used Frequentist one-sample t-tests to evaluate whether the accuracy and RT interferences scores on the modified “my priest” psychology and physicality items were different from zero. Results revealed that interference scores were not significantly different from zero for physicality RT, *t* (68) < 1.0, *p* > .250, and for psychology accuracy, *t* (68) < 1.0, *p* > .250, and RT, *t* (68) = 1.21, *p* = .229. Results for physicality accuracy were significantly different from zero, *t* (68) = 2.58, *p* = .012, *d* = -0.31, 95% CI = [-0.55, -0.07], however the effect was in the opposite direction to that in Exp. 1 and 3.

We next performed the following additional checks. First, we confirmed that interference scores on the unmodified “God” psychology items were different from zero: accuracy: *t* (68) = 3.80, *p* < .001, *d* = 0.46, 95% CI = [0.21, 0.70]; RT: *t* (68) = 2.98, *p* = .004, *d* = 0.36, 95% CI = [0.11, 0.60]. We thereby showed that the null effects for the “my priest” items were not due to a general failure to replicate the interference effects in this experiment. Second, we confirmed that interference scores on the same subset of items that were modified in this experiment (“my priest”) were for the most part different from zero in Exp. 1 and 3 where they are unmodified (i.e., where they read “God”): Exp. 1, physicality: accuracy: *t* (126) = 4.85, *p* < .001, *d* = 0.43, 95% CI = [0.25, 0.61]; RT: *t* (126) = 5.91, *p* < .001, *d* = 0.52, 95% CI = [0.34, 0.71]; psychology: accuracy: *t* (126) = 4.51, *p* < .001, *d* = 0.40, 95% CI = [0.22, 0.58]; RT: *t* (126) = 0.59, *p* = .56, *d* = 0.05, 95% CI = [-0.12, 0.23]; Exp. 3, physicality: accuracy: *t* (67) = 4.10, *p* < .001, *d* = 0.50, 95% CI = [0.24, 0.75]; RT: *t* (67) = 5.54, *p* < .001, *d* = 0.67, 95% CI = [0.41, 0.93]; psychology: accuracy: *t* (67) = 3.85, *p* < .001, *d* = 0.47, 95% CI = [0.22, 0.72], RT: *t* (68) = 0.17, *p* = .867, *d* = 0.02, 95% CI = [-0.22, 0.26]. We thereby showed that the effects for the “my priest” items observed in this experiment were not due to a failure to replicate the interference effects in that subset of items.

**3. Multiverse analyses**

**3.1. Exp. 1**





**3.2. Exp. 3**





**3.3. Exp. 4**





**4. Full List of Items**











