Effects of Mindfulness on Conceptual Modeling Performance: a Series of Experiments

Supplemental Material: Detailed Analysis of Results of MIND#3 and Additional Analyses of the pooled data

Beatriz Bernárdez, Amador Durán, José A. Parejo, Natalia Juristo, Antonio Ruiz–Cortés

Abstract—This supplemental material contains the detailed analysis of MIND#3, the second internal replication whose details have not been previously published elsewhere, and some additional analyses of the pooled data not included in the main article.

1 ANALYSIS OF RESULTS OF MIND#3

This section contains a detailed analysis of the second internal replication, MIND#3.

1.1 Descriptive Statistics of MIND#3

The descriptive statistics of MIND#3 are shown in Table 3, including measures of central tendency (mean and median), and variability (standard deviation, minimum, maximum and range values). The distributions of the response variables are depicted as box plots in Figure 1. The only outlier in the sample for conceptual modeling efficiency in both the pre and post–treatment exercises was subject 38. After carefully scrutinizing his data and finding no anomalies, he was considered as a genuine outlier, i.e. the subject was simply faster than the others in the control group.

1.2 Hypotheses Tests in MIND#3

Although the probability of a random group assignment to produce a decompensated distribution is very low [1], the differences between the scores of the experimental and control groups in the pre–treatment exercise were examined as a double–check using a one–way ANOVA as a homogeneity test for both response variables, as recommended in [1]. Since the null hypotheses were not rejected (p–values for effectiveness and efficiency were 0.912 and 0.421 respectively), there was no evidence of significant differences between groups before the administration of treatment.

Once homogeneity was verified, assumptions for parametric statistical tests were also checked. Normality and *homoscedasticity* (homogeneity of variances) were verified using a Shapiro–Wilk and a Levene test respectively (see

Table 1
Shapiro–Wilk normality test results for MIND#3 (including subject 38)

Dependent variable	Group	Time	Sig.
Conceptual	mindfulness	pre–treatment	0.998
Modeling	control	pre–treatment	0.947
Effectiveness	mindfulness	post–treatment	0.621
	control	post–treatment	0.117
Conceptual	mindfulness	pre–treatment	0.131
Modeling	control	pre–treatment	0.001 *
Efficiency	mindfulness	post–treatment	0.614
	control	pre–treatment	0.003 *

Table 2
Levene test results (including subject 38)

Dependent variable	F	Sig.
CM Effectiveness	0.602	0.440
CM Efficiency	7.392	0.008 *

Tables 1 and 2). All the *time* × *group* combinations for both response variables passed the Shapiro–Wilk tests except those of the control group for efficiency, which did not pass the Levene test either (marked with an asterisk in Tables 1 and 2) because of subject 38, who was faster than the rest of students in the control group, as previously discussed. However, the results of the same tests excluding that subject passed the tests, so finally a 2 (*group*) × 2 (*time*) mixed–model ANOVA was performed for each response variable because of the robustness of the test, as stated in [2].

In a similar way to MIND#1–2, the results of the 2×2 mixed–model ANOVA in Tables 4 and 5 show a *small* but non–significant effect of mindfulness on conceptual modeling effectiveness ($F(1,43) = 2.614$, $p = 0.113$), and a *large* highly significant effect on efficiency ($F(1,43) = 61.602$, $p < 0.01$), as described by the *group*time* interactions.

- *B. Bernárdez, A. Durán, J. A. Parejo and A. Ruiz–Cortés are with the E.T.S.I. Informática, Universidad de Sevilla, Spain. E-mail: beat@us.es, amador@us.es, japarejo@us.es, aruiz@us.es*
- *N. Juristo is with the E.T.S.I. Informática, Universidad Politécnica de Madrid, Spain. E-mail: natalia@fi.upm.es*

Manuscript received Month day, 2018; revised Month day, 2018.

Table 3
Descriptive statistics of MIND#3

	Conceptual Modeling Effectiveness				Conceptual Modeling Efficiency			
	Control group		Experimental group		Control group		Experimental group	
	Pre–treatment	Post–treatment	Pre–treatment	Post–treatment	Pre–treatment	Post–treatment	Pre–treatment	Post–treatment
n	22	22	23	23	22	22	23	23
mean	0,511	0,409	0,507	0,455	0,295	0,261	0,316	0,434
sd	0,123	0,102	0,127	0,087	0,076	0,085	0,091	0,100
median	0,530	0,460	0,524	0,457	0,285	0,241	0,295	0,432
min	0,238	0,260	0,238	0,300	0,213	0,158	0,185	0,273
max	0,762	0,730	0,762	0,600	0,552	0,547	0,509	0,617
range	0,524	0,470	0,524	0,300	0,339	0,389	0,324	0,344

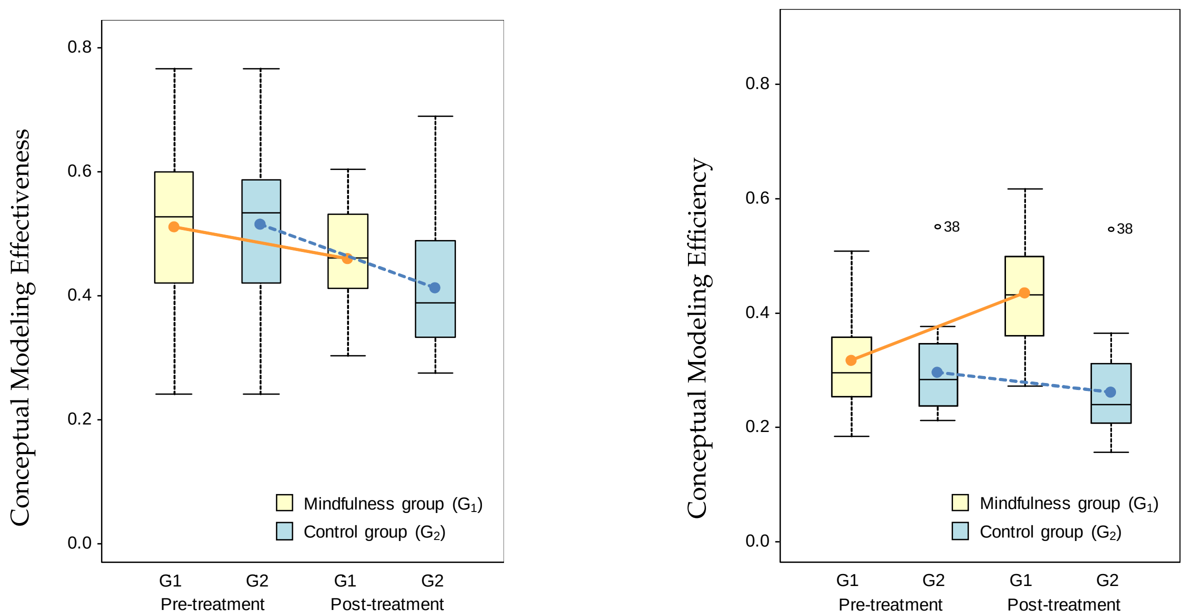

Figure 1. Box Plots of Response Variables in MIND#3

Table 4
2 (*group*)×2 (*time*) Mixed–model ANOVA of
Conceptual Modeling Effectiveness in MIND#3

Source of variation	Sum of Squares	DF	Mean Square	F	Sig.	η_p^2
time	0.135	1	0.135	25.393	0.000	0.371
*group*time*	0.014	1	0.014	2.614	0.113	0.057
Error	0.228	43	0.005			

Table 5
2 (*group*)×2 (*time*) Mixed–model ANOVA of
Conceptual Modeling Efficiency in MIND#3

Source of variation	Sum of Squares	DF	Mean Square	F	Sig.	η_p^2
time	0.041	1	0.041	19.328	0.000	0.310
*group*time*	0.132	1	0.132	61.602	0.000	0.590
Error	0.091	43	0.002			

Table 6
ANCOVA of Conceptual Modeling Effectiveness in Pooled Data with *postscore* as the response variable and *prescore* as a covariate

Source of variation	Df	Sum of Squares	Mean Squares	F	Sig.	η_p^2
prescore	0,035	1	0,035	1,773	0,185	0,014
group	0,085	1	0,085	4,342	0,039	0,033
Error	2,478	127	0,020			

Table 7
ANCOVA of Conceptual Modeling Efficiency in Pooled Data with *postscore* as the response variable and *prescore* as a covariate

Source of variation	Df	Sum of Squares	Mean Squares	F	Sig.	η_p^2
prescore	0,060	1	0,060	4,616	0,034	0,035
group	0,624	1	0,624	48,272	0,000	0,275
Error	1,642	127	0,013			

Table 8
2 (*group*)×2 (*post-task*) Mixed–model ANCOVA of Conceptual Modeling Effectiveness in Pooled Data with *postscore* as the response variable and *prescore* as a covariate

Source of variation	Sum of Squares	DF	Mean Square	F	Sig.	η_p^2
prescore	0.243	1	0.243	16.369	0.000	0.116
group	0.085	1	0.085	5.739	0.018	0.044
post-task	0.623	1	0.623	42.014	0.000	0.252
*group***post-task*	0.001	1	0.001	0.092	0.762	0.001
Error	1.854	125	0.015			

Table 9
2 (*group*)×2 (*post-task*) Mixed–model ANCOVA of Conceptual Modeling Efficiency in Pooled Data with *postscore* as the response variable and *prescore* as a covariate

Source of variation	Sum of Squares	DF	Mean Square	F	Sig.	η_p^2
prescore	0.125	1	0.125	10.212	0.002	0.076
group	0.607	1	0.607	49.589	0.000	0.284
post-task	0.098	1	0.098	8.003	0.005	0.060
*group***post-task*	0.014	1	0.014	1.131	0.290	0.009
Error	2.478	125	0.020			

Table 10
2 (*group*)×2 (*time*) Mixed–model ANCOVA of Conceptual Modeling Effectiveness in Pooled Data with *task* as a covariate

Source of variation	Df	Sum of Squares	Mean Squares	F	Sig.	η_p^2
task	1	1.055	1.055	102.237	0.000	0.446
time	1	0.088	0.088	8.511	0.004	0.063
*group***time*	1	0.074	0.074	7.126	0.009	0.053
Error	127	1.310	0.010			

Table 11
2 (*group*)×2 (*time*) Mixed–model ANCOVA of Conceptual Modeling Efficiency in Pooled Data with *task* as a covariate

Source of variation	Df	Sum of Squares	Mean Squares	F	Sig.	η_p^2
task	1	0.537	0.537	71.941	0.000	0.362
time	1	0.630	0.630	84.485	0.000	0.399
*group***time*	1	0.274	0.274	36.748	0.000	0.224
Error	127	0.947	0.007			

2 Additional Analyses of the Pooled Data

For the sake of completeness, and aiming at providing further insights and confidence on the effect of mindfulness on conceptual modeling, some additional analyses of the pooled data of MIND#1–3 are provided in this section.

The first additional analysis was an ANCOVA using the post–treatment score as the dependent variable and the pre–treatment score (*prescore*) as a covariate, which is a usual alternative to mixed–model ANOVAs for pre–post designs [3], [4]. This analysis assesses the differences in the post–treatment means after accounting for the pre–treatment values, focusing the conclusions more on the effect of the treatment than in the temporal evolution of the phenomenon under study, as a mixed–model ANOVA does.

In our case, the results of applying this analysis to the two response variables are shown in Tables 6 and 7, which show very similar results to the 2×2 mixed–model ANOVAs in the main article, i.e. a *small* significant effect for effectiveness ($F(1,127) = 4.342$, $p < 0.05$), and a *large* highly significant effect for efficiency ($F(1,127) = 48.272$, $p < 0.01$), as described by the *group* factor.

Considering the obtained results about the effect of the order in which the conceptual modeling exercises were performed in the main article, i.e. that it had a significant impact but its interaction with the treatment was not significant, we decided to study its effect by adding it as a factor to the previously discussed ANCOVAs as we did for the 2×2 mixed–model ANOVAs in the main article. Note that in this case, where the post–treatment score (*postscore*) is the response variable, the order is described by the post–treatment exercise, represented by the *post-task* factor with ERASMUS and EODP levels in Tables 8 and 9.

As in the main article, introducing a factor related with the order as *post-task* yields not only lower p–values for effectiveness ($F(1,125) = 5.739$, $p < 0.05$) and efficiency ($F(1,125) = 49.589$, $p < 0.01$), but also bigger effect sizes ($\eta_p^2 = 0.044$ and $\eta_p^2 = 0.284$ respectively), as described by the *group* factor. Note also that, although the effect of the post–task is significant for both response variables, its interaction with the treatment is not significant for either effectiveness ($F(1,125) = 0.092$, $p = 0.762$) or efficiency ($F(1,125) = 1.131$, $p = 0.290$), as described by the *group***post-task* interaction.

Finally, as a double–check for the mixed–model ANCOVAs with the order as a covariate reported in the main article, a mixed–model ANCOVA with the task as a covariate were also conducted for each response variable, as shown in Tables 10 and 11.

As expected, both tests show exactly the same values for the effect of the treatment than in the main article, i.e. a *small* highly significant effect for effectiveness (F(1,127) = 7.126, $p < 0.01$), and a *large* highly significant effect for efficiency (F(1,127) = 36.748, $p < 0.01$), as described by the *group*time* interactions.

References

[1] D. T. Campbell and S. Julian, *Experimental and Quasi–Experimental Designs for Research*. United States: Wadsworth, 1963.
[2] G. Norman, "Likert scales, levels of measurement and the laws of statistics," *Advances in Health Sciences Education*, vol. 15, no. 5, pp. 625–632, 2010.
[3] E. L. Fancher, "Comparison of methods of analysis for pretest and posttest data," Master's thesis, University of Georgia, 2013.
[4] A. Rutherford, *Introducing ANOVA and ANCOVA: a GLM approach*. Sage, 2001.