Corrigendum

aps ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Psychological Science
2018, Vol. 29(9) 1565–1568

DOI: 10.1177/0956797618781773
www.psychologicalscience.org/PS
SAGE

Corrigendum: The Pen Is Mightier Than the Keyboard: Advantages of Longhand Over Laptop Note Taking

Original article: Mueller, P. A., & Oppenheimer, D. M. (2014). The pen is mightier than the keyboard: Advantages of longhand over laptop note taking. *Psychological Science*, *25*, 1159–1168. doi:10.1177/0956797614524581

This Corrigendum corrects five categories of errors in the original article:

- The *z* scores reported in the data files uploaded to the Open Science Framework (OSF) were calculated prior to having to remove 1 participant in Study 1, 2 participants in Study 2, and 9 participants in Study 3 for various reasons (which were reported in the original article). We had removed those participants from the data files on OSF but had failed to recalculate the *z* scores without those participants included.
- In the original article, the *z* scores from Study 1 were calculated using an index score across lectures (a perfect score would be 1 point per question; 10 points total), whereas the *z* scores in Studies 2 and 3 were calculated using the raw data (in which different questions had different point values). The article erroneously indicated that the index-score approach was used for all three studies. We will now report all results using index scores; the pattern of results is the same for both measures.
- Moreover, the data in Study 2 were *z*-scored within lectures, which limited the inferential power of the analyses. We will now report results from Study 2 using *z* scores across lectures. The patterns of results do not appreciably change.
- The article mistakenly reported the degrees of freedom for error for the interaction of condition and lecture in the first two studies (55 and 89 for Study 1 and Study 2, respectively) instead of the degrees of freedom for error for condition (4.01 and 4.09 for Study 1 and Study 2, respectively, due to the use of mixed and random-effects analyses).
- Some of the effect sizes were reported as η_p^2 values when they were actually η^2 values. The effect sizes reported in this Corrigendum (and which will be corrected in the article) are the η_p^2 values, as presented in SPSS.

New data files with the corrected *z* scores, as well as SPSS syntax for the corrected analyses, have been uploaded to OSF (https://osf.io/crsiz). A file with the syntax for the originally uploaded files with the *z*-scoring errors has also been included (see the associated Wiki for detailed descriptions). Additionally, because of the large number of experimenter degrees of freedom in these analyses (e.g., whether we used index or raw scores), to demonstrate that the effects are robust beyond specific analysis decisions, we have posted a series of analyses showing how the results change across different analysis strategies. We regret the errors and inconsistencies in the original manuscript but are heartened that regardless of how the data are analyzed, the results remain qualitatively the same.

Results that are now being corrected in the article are as follows.

Study 1

On page 1161, in the Results section of Study 1, the values in the second half of the first paragraph will be changed as follows (the text and overall conclusions remain the same):

> On factual-recall questions, participants performed equally well across conditions (laptop: $M = -0.006$, $SD = 1.00$; longhand: $M = 0.05$, $SD = 1.01$), $F(1, 4.01) = 0.046$, $p = .841$. However, on conceptual-application questions, laptop participants performed significantly worse ($M = -0.178$, $SD = 0.900$) than longhand participants ($M = 0.162$, $SD = 1.07$), $F(1, 4.09) = 8.05$, $p = .046$, $\eta_p^2 = .66$ (see Fig. 1).[5] Which lecture participants saw also affected performance on conceptual-application

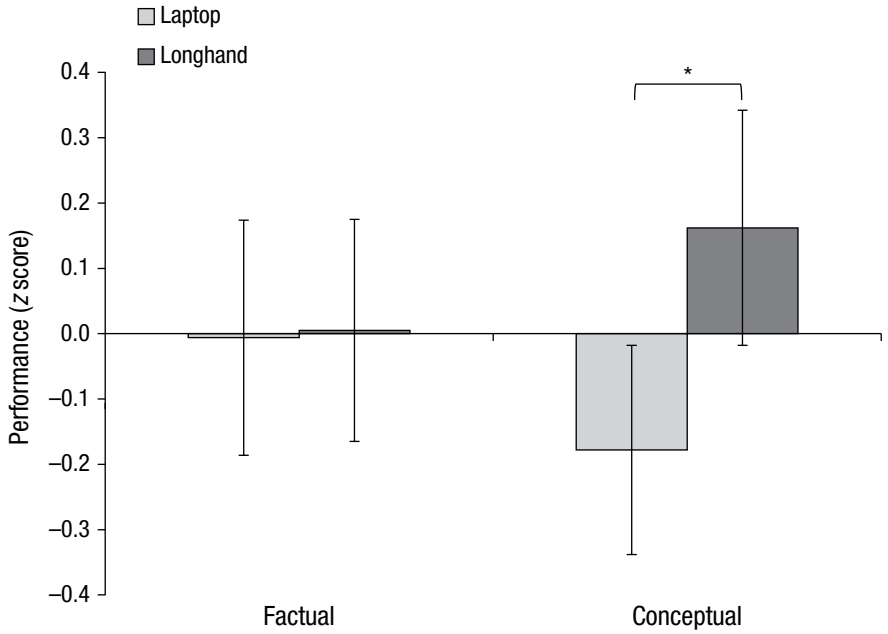

Fig. 1. Mean performance on factual-recall and conceptual-application questions as a function of note-taking condition (Study 1). Performance was converted to an index score where each question was worth 1 point, and the totals were *z*-scored across lectures. The asterisk indicates a significant difference between conditions ($p < .05$). Error bars indicate standard errors of the mean.

questions, $F(4, 4) = 7.11$, $p = .042$, $\eta_p^2 = .88$; however, there was no significant interaction between lecture and note-taking medium, $F(4, 55) = 0.259$, $p = .90$.

Figure 1 will also be corrected to show the new means, and the caption will be changed as indicated.

In the original article, we did not report the null effect of performance across lectures on factual questions. To fill that omission, we now report the following at the end of the first paragraph on p. 1161: "There was no significant difference in performance on factual questions across lectures, $F(4, 4) = 1.57$, $p = .33$."

Study 2

On page 1162, the first paragraph of the "Laptop versus longhand performance" subsection will be replaced with the following (the overall conclusions remain the same):

> Responses were scored by raters blind to condition. Replicating our original finding, results showed that on conceptual-application questions, longhand participants performed better (*z*-score $M = 0.24$, $SD = 1.11$) than laptop-nonintervention participants (*z*-score $M = -0.17$, $SD = 0.88$), $F(1, 6.60) = 19.65$, $p = .003$, $\eta_p^2 = .75$. Scores for laptop-intervention participants (*z*-score $M = -0.11$, $SD = 1.02$) did not significantly differ from those for either laptop-nonintervention ($p = .91$) or longhand ($p = .29$) participants. The pattern of data for factual questions was similar, though the differences were not significant (longhand: *z*-score $M = 0.025$, $SD = 0.97$; laptop intervention: *z*-score $M = 0.063$, $SD = 1.05$; laptop nonintervention: *z*-score $M = -0.089$, $SD = 0.99$), $F(1, 4.54) = 4.08$, $p = .11$ (see Fig. 4).[8] There was a significant difference in conceptual performance across lectures, $F(4, 4) = 19.87$, $p = .007$, $\eta_p^2 = .95$, but the interaction was not significant, $F(4, 89) = 0.138$, $p = .97$. There was a significant difference in factual performance across lectures, $F(4, 4) = 14.59$, $p = .012$, $\eta_p^2 = .94$, but the interaction was not significant, $F(4, 89) = 0.439$, $p = .66$.

Figure 4 will also be corrected to show the new means, and the caption will be changed as indicated.

In addition, a new endnote (Note 9) will be added at the end of the Results section for Study 2 (p. 1163). The text of this note will read as follows:

> Estimating effect sizes in mixed models is problematic (for more details, see, e.g., Judd, Westfall, & Kenny, 2017). Thus, while the η_p^2s reported above for Studies 1 and 2 (ranging from .66 to .95) are what SPSS reports, more informative estimates of effect sizes would be much smaller, ranging from .03 to .11.

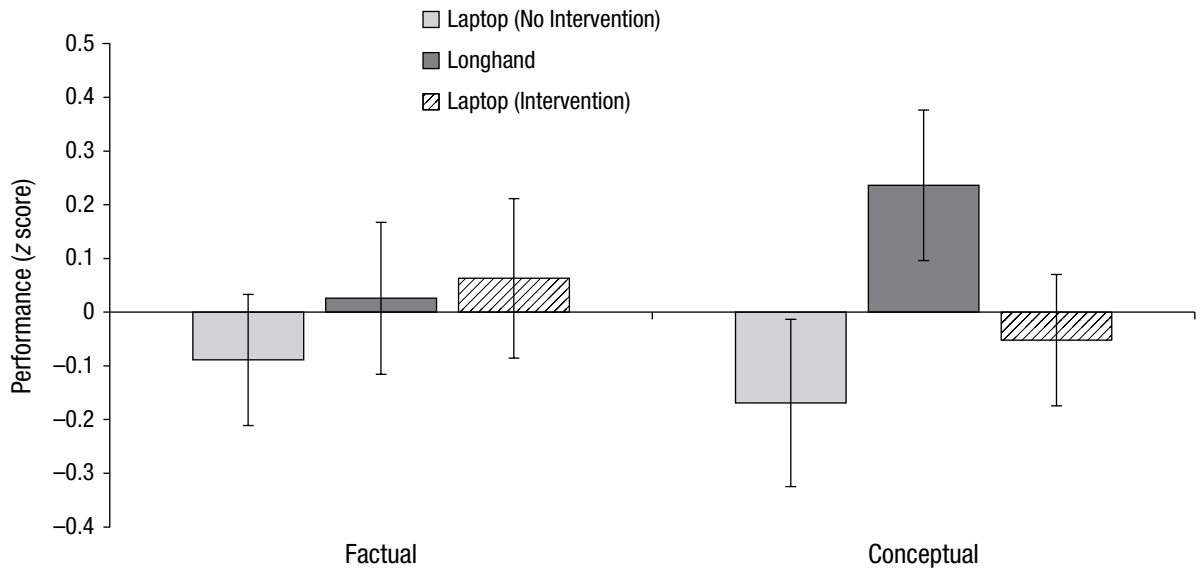

Fig. 4. Mean performance on factual-recall and conceptual-application questions as a function of note-taking condition (Study 2). Performance was converted to an index score where each question was worth 1 point, and the totals were *z*-scored across lectures. Error bars indicate standard errors of the mean.

Study 3

Results for Study 3 were calculated using *z* scores of raw rather than index scores. The overall conclusions stand, but the two paragraphs in the "Laptop versus longhand performance" subsection (p. 1164) will be changed as follows:

> Across all question types, there were no main effects of note-taking medium or opportunity to study. However, there was a significant interaction between these two variables, $F(1, 105) = 5.62$, $p = .02$, $\eta_p^2 = .05$. Participants who took longhand notes and were able to study them performed significantly better (*z*-score $M = 0.45$) than participants in any of the other conditions (*z*-score *M*s = −0.25, −0.02, −0.20), $t(105) = 2.86$, $p = .005$, $d = 0.56$ (see Fig. 5).
>
> Collapsing questions about facts and seductive details into a general measure of "factual" performance, we found a significant main effect of note-taking medium, $F(1, 105) = 4.05$, $p = .047$, $\eta_p^2 = .037$, and of opportunity to study, $F(1, 105) = 11.49$, $p = .001$, $\eta_p^2 = .09$, but this was qualified by a significant interaction, $F(1, 105) = 5.45$, $p = .021$, $\eta_p^2 = .05$. Again, participants in the longhand-study condition (*z*-score $M = 0.68$) outperformed the other participants (*z*-score *M*s = −0.09, −0.28, −0.34), $t(105) = 4.50$, $p < .001$, $d = 0.88$. Among students who had the opportunity to study, longhand note takers did significantly better than laptop note takers, $t(53) = 2.77$, $p = .008$.
>
> Collapsing performance on conceptual, inferential, and application questions into a general "conceptual" measure revealed no significant main effects, but again there was a significant interaction between note-taking medium and studying. There was a significant interaction for conceptual questions, $F(1, 105) = 4.35$, $p = .04$, $\eta_p^2 = .04$. Among students who had the opportunity to study, longhand note takers did significantly better than laptop note takers, $t(53) = 2.32$, $p = .024$, $d = 0.64$ (for raw means, see Table 2).

The following sentence will be added for clarification to the note to Table 2: "Although raw scores are given here, *z* scores of the index scores were used in the analysis." Figure 5 will also be corrected to show the new means, and the caption will be changed as indicated.

A new endnote (Note 10) will be added at the end of the Results section for Study 3 (p. 1166). The text of this note will read as follows:

> Participants who took laptop or longhand notes but who did not have the opportunity to study did not score significantly differently on either factual, $t(52) = 0.26$, $p = .795$, $d = 0.07$, or conceptual, $t(52) = 0.77$, $p = .442$, $d = 0.21$, questions.

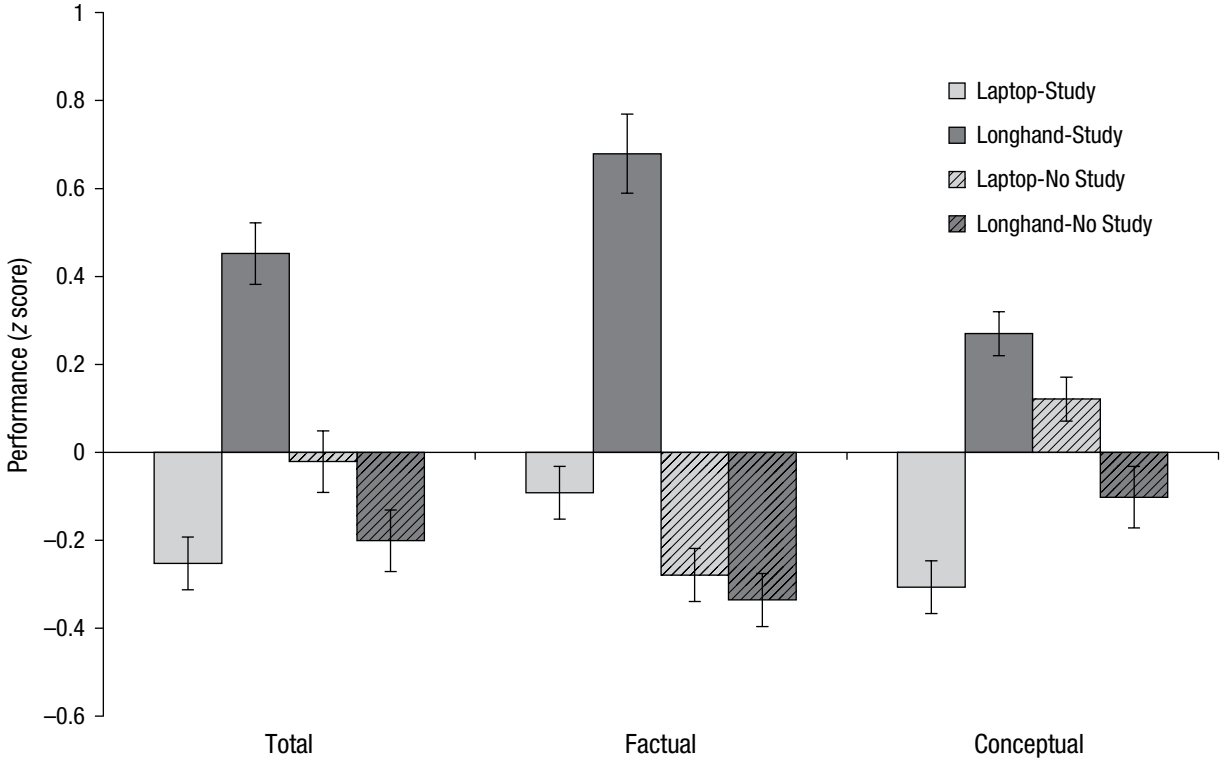

Fig. 5. Mean performance on factual-recall and conceptual-application questions as a function of note-taking condition and opportunity to study (Study 3). Performance was converted to an index score where each question was worth 1 point, and the totals were *z*-scored across lectures. Error bars indicate standard errors of the mean.

References

The following reference will be added to the References section:

Judd, C. M., Westfall, J., & Kenny, D. A. (2017). Experiments with more than one random factor: Designs, analytic models, and statistical power. *Annual Review of Psychology*, *68*, 601–625.

Corresponding Author

The Corresponding Author will be changed to Daniel M. Oppenheimer (p. 1159). The new contact information is as follows: Carnegie Mellon University, Department of Social and Decision Sciences and Department of Psychology, Porter Hall 208, 5000 Forbes Ave., Pittsburgh, PA 15213. E-mail: oppenheimer@cmu.edu