REPLY

Supersizing Effect Sizes Raises Concerns: A Reply to Zimmerman (2013)

AQ: au

M. Brent Donnellan
Michigan State University

Christopher J. Ferguson
Stetson University

AQ: 1

Zimmerman (2013) suggested that our reanalysis adds little to the scientific literature. We disagree. We clarify our motivations and explain how further analyses based on his suggestion for age do not change our conclusions. Moreover, we believe the nascent experimental literature is more in line with our interpretations than Zimmerman's. We continue to believe that scientists should be cautious when describing effect sizes.

Keywords: mass media, television, language development, cognition, child development

We concluded that the Zimmerman, Christakis, and Meltzoff (2007) data set does not provide robust support for the conclusion of a "large" negative association between baby video exposure and language development. Zimmerman (2013) offered an interesting reaction to our re-analysis but did little to alter our conclusion. He criticized our motivations, literature review, and approach, but did little to counter our primary arguments. Zimmerman could have explained how the original data set supported his conclusions, but he focused on other topics. We briefly respond to several issues raised in his comment.

What Were Our Motivations?

Zimmerman (2013) suggested that we are biased and "dissatisfied" with the results reported in Zimmerman et al. (2007). This seems unnecessarily inflammatory as we have no vested interest in the original results per se. We were concerned about the limited reporting in the 2007 report (e.g., zero-order correlations were not described) and confused as to why they did not report results for both the expressive and receptive language scores for the younger cohort. These considerations made it difficult to thoroughly evaluate their results. Zimmerman further suggested that our concerns about publication bias were a major motivator of our re-analysis. However, we stated that our re-analysis was primarily motivated AQ: 2 by "concerns over research degrees of freedom" (p. XXX). In particular, we were puzzled to learn that a −.07 correlation transmogrified into a "large" effect in the context of a multivariate analysis.

In short, we were motivated by curiosity. Zimmerman et al. (2007) started their Discussion with the following restatement of their results:

This analysis reveals a large negative association between viewing of baby DVDs/videos and vocabulary acquisition in children age 8 to 16 months. The 17-point difference associated in the analysis with each hour of baby DVD/video watching corresponds to a difference of about 6 to 8 words for a typical child out of the 90 included on the CDI. (p. 366)[1]

Fn1

This interpretation was based on the −16.99 unstandardized coefficient for the baby video exposure variable from their multiple regression model. We can translate the −.07 correlation into an unstandardized regression coefficient by taking the standard deviation for the language scores reported in their article, dividing it by the standard deviation for the exposure variable, and multiplying it by the correlation of −.07. This yields a coefficient of −5.44. Thus, the inclusion of covariates tripled a −5.44 coefficient to −16.99 in the original report, and this increase in effect size seemed worthy of investigation.

What Language Dependent Variable Should Have Been Used?

Zimmerman argued that it is appropriate to use age-normalized language scores as opposed to our approach of using raw scores with controls for age and gender. One of our reservations about the original Zimmerman approach was that the normative sample was not created with scientific sampling methods. He did not share our reservations but suggested an alternative approach using complex polynomial terms for age. Regardless, the decision to use age-normalized scores as opposed to controlling for age in some fashion is a choice. Readers of the original article could have been notified to the extent that results differed across different methods for handling the dependent variable. This kind of flexibility is an example of our concerns about the impact of researcher degrees of freedom.

Correspondence concerning this article should be addressed to Christopher J. Ferguson, Department of Psychology, Stetson University, 421 N. Woodland Blvd., DeLand, FL 32729. E-mail: cjferguson1111@aol.com

[1] Zimmerman argued that we were misleading in how we discussed the Zimmerman et al. (2007) article. We disagree and hope interested readers will read their article to evaluate how we described their results in our article. The phrase "large" is used at least twice in the original article to describe the focal effect size.

We conducted some preliminary analyses to evaluate Zimmerman's suggestion. We regressed receptive language scores for the younger sample on centered age, centered age^2, centered age^3, gender, and raw scores for exposure to baby videos. There was no evidence of a statistical effect for exposure to baby videos, $B = -0.008$; $\beta = -.008$, $p = .848$; $R^2 = .371$. As a point of comparison, the standardized coefficient was $-.01$ in our Table 2. We also used expressive language as the outcome and found similar results for the baby video exposure variable to those reported in our Table 2, $B = 0.035$; $\beta = .064$, $p = .153$; $R^2 = .248$. Altogether, this suggests that controlling for age^2 and age^3 terms would not have altered our basic conclusions. Indeed, our inspection of the data did not suggest a strong nonlinear association between age and language scores within the limited age window for the younger cohort. Again, these alternative analyses do not support a conclusion of a "large" negative effect for baby video viewing.

Should We Have Transformed Media Exposure Variables?

AQ: 3

Zimmerman critiqued our use of square-root transformations for media exposure variables stating that "any transformation . . . replace[s] observed data points with different data points and, in this sense, introduce[s] measurement error into a key variable" (p. XXX). A square-root transformation is a re-expression of the original scores using a simple mathematical operation. Measurement error (at least as we understand the concept) was not introduced in this process. If a parent reported that her or his child watched 9 min of baby videos a day, there is no reason to believe that re-expressing the "9" as a "3" by taking the square-root introduces any more or less precision to that measurement of exposure. The media exposure variables were skewed, and it is reasonable to evaluate how methods to address this issue impact conclusions. Accordingly, we consistently reported results using both raw scores and square-root transformed variables in our article. Readers can judge for themselves about the importance of the transformations. The different variables produced generally similar effect sizes, and none were large.

What Do Experimental Studies Indicate?

Zimmerman suggested that we presented a one-sided literature review in our brief report. In our defense, we cited a 2009 review article that concluded there was a need for more research. This seems like an uncontroversial conclusion. Zimmerman pointed to two recent experimental studies on this topic; however, both returned null results for their experiments (e.g., Richert, Robb, Fender, & Wartella, 2010; Robb, Richert, & Wartella, 2009). Richert et al. (2010) noted that "exposure to the DVD was unrelated to our measures of general language learning, providing no evidence that exposure to this DVD over 6 weeks either helped or hindered children's general language learning" (p. 436). This result seems more consistent with our conclusions than Zimmerman's. The Richert et al. study had 44 children in each of two conditions (DVD vs. no DVD), and a cursory power calculation indicates that their study was adequately powered to detect a "large" effect size (i.e., power over 90% to detect a standardized mean difference of .80 or more). Robb et al. (2009) also yielded null experimental results although it included fewer participants and thus had less power. Unless we overlooked critical findings, the nascent experimental literature does not support the conclusion of large media effects.

Conclusion

We hope that readers are not turned off by the tone of the Zimmerman exchange and simply tune out. There are important issues at stake. First and foremost, we are concerned about the impact of researcher degrees of freedom on results, and we think the field benefits by grappling with these issues. There are no "magic bullet" solutions, but greater awareness itself is valuable. At best, the field can continue to promote constructive criticism of research, increase transparency in the form of posting raw data, and facilitate a meaningful post-peer review process. Our article was intended to further post-peer review, and we believe it has been successful in this regard. Second, we believe that researchers need to exercise caution when interpreting their findings. Descriptors like "large" should be used with caution and only when strongly justified by data. This kind of restraint is especially important in the context of engagement with the broader public. There are risks to the field in terms of diminished credibility when claims about large effect sizes are not well supported.

References

Ferguson, C. J., & Donnellan, M. B. (2013, July 13). Is the association between children's baby video viewing and poor language development robust? A reanalysis of Zimmerman, Christakis, and Meltzoff. (2007). *Developmental Psychology*. Advance online publication. doi: 10.1037/a0033628

Richert, R. A., Robb, M. B., Fender, J. G., & Wartella, E. (2010). Word learning from baby videos. *Archives of Pediatrics & Adolescent Medicine, 5,* 432–437. doi:10.1001/archpediatrics.2010.24

Robb, M. B., Richert, R. A., & Wartella, E. A. (2009). Just a talking book? Word learning from watching baby videos. *British Journal of Developmental Psychology, 27,* 27–45. doi:10.1348/026151008X320156

Zimmerman, F. J. (2014). Where's the beef? A commentary on Ferguson and Donnellan. *Developmental Psychology, 50,* XXX–XXX.   **AQ: 4**

Zimmerman, F. J., Christakis, D. A., & Meltzoff, A. N. (2007). Associations between media viewing and language development in children under age 2 years. *The Journal of Pediatrics, 151,* 364–368. doi: 10.1016/j.jpeds.2007.04.071