Perspectives on Psychological Science
2016, Vol. 11(3) 408–414

DOI: 10.1177/1745691616635599
pps.sagepub.com

Angry Birds, Angry Children, and Angry Meta-Analysts: A Reanalysis

Luis Furuya-Kanamori and Suhail A. R. Doi
Research School of Population Health, Australian National University, Canberra, Australia

Abstract

Ferguson's (2015a) meta-analysis assessed a very important and controversial topic about children's mental health and video games. In response to the concerns raised by researchers about the appropriateness of the meta-analytical methods used by Ferguson; we decided to reanalyze the data and discuss two major misconceptions about meta-analysis. We argue that partial correlations can (and should) be meta-analyzed instead of zero-order bivariate correlations if the predictors included in the partial correlation represent a similar construct. We also discuss the fallacy by which the conventional meta-analytical model assumes that the studies' effect sizes came into being according to the same random effect construct used by the analysis. Our replication results using partial correlations, standardized (valid and reliable) outcomes, and an improved meta-analytical model (that does not assume a random effect is the mechanism of data generation) confirmed the main results of Ferguson's meta-analysis. There was a significant yet very small effect on aggressive behavior of exposure to both general, $r_p = 0.062$, 95% CI [0.012, 0.112], and violent, $r_p = 0.055$, 95% CI [0.019, 0.091], video games. A very small effect was seen on reduced prosocial behavior, but this was only in the general video game exposure category, $r_p = 0.072$, 95% CI [0.045, 0.100].

Keywords

video games, aggression, meta-analysis, partial correlation, publication bias

We read with great interest Ferguson's "Angry Birds" meta-analysis (Ferguson, 2015a) and the subsequent commentaries (Boxer, Groves, & Docherty, 2015; Ferguson, 2015b; Gentile, 2015; Markey, 2015; Rothstein & Bushman, 2015; Valkenburg, 2015), which covered the effect of video games on mental health and meta-analytical methods—two topics that are currently fiercely debated in the scientific community. Given the debate generated by several researchers in their comments (Boxer et al., 2015; Gentile, 2015; Markey, 2015; Rothstein & Bushman, 2015; Valkenburg, 2015) regarding the methodology used by Ferguson, we decided to reanalyze the data and address some misconceptions regarding meta-analytical methods.

Reanalysis of the Data

We reexamined the effect of videos games on childhood and teenage aggression. Our analyses used Ferguson's data, albeit restricted to those studies that applied standardized outcome measures in order to avoid the inclusion of studies in which the scale has not yet been validated or tested for its reliability. Two meta-analytical models, the inverse variance heterogeneity (IVhet; Doi, Barendregt, Khan, Thalib, & Williams, 2015) and the random effects (RE; DerSimonian & Laird, 1986) models, were used to pool the effect size estimates for video game exposure and aggressive behavior, reduced prosocial behavior, reduced academic performance, depressive symptoms and attention deficit symptoms. Subgroup analyses were conducted to determine if exposure to different video games (violent vs. general) had a different effect on children and adolescents. Publication bias was examined through visual inspection of the funnel and Doi plots as well as quantitative assessment of asymmetry through the LFK index (Barendregt & Doi, 2015). The Doi plot uses a rank-based measure (Z score) of precision (instead of the standard error) and plots it against the effect

Corresponding Author:
Suhail A. R. Doi, Research School of Population Health, Australian National University, 62 Mills Road, Acton, ACT 2601, Australia
E-mail: sardoi@gmx.net

size. Asymmetrical Doi plots suggest presence of publication bias. The LFK index quantifies the extent of Doi plot asymmetry by averaging half of the sum of the Z score plus the normalized effect size across the meta-analysis. The meta-analyses were conducted using MetaXL version 3.0 (EpiGear; Sunrise Beach; Australia; http://www.epigear.com).

After applying the exclusions for unstandardized outcome measures and bivariate correlations, we were left with 35, 12, 12, 16, and 6 studies for aggressive behavior, reduced prosocial behavior, reduced academic performance, depressive symptoms, and attention deficit symptoms, respectively. The results showed that exposure to video games were associated (though with a very small effect size) with aggressive behavior overall, $r_p = 0.059$, 95% CI [0.028, 0.090], and across both the general, $r_p = 0.062$, 95% CI [0.012, 0.112], and violent, $r_p = 0.055$, 95% CI [0.019, 0.091], subgroups of video game exposure. The only other statistically significant association, also with a very small effect size, was found with reduced prosocial behavior, but this was (paradoxically) only associated with general video game exposure, $r_p = 0.072$, 95% CI [0.045, 0.100], and not with violent video game exposure. No significant associations were found for the other outcomes measured (Table 1). No asymmetry was identified across all studies for aggressive behavior and reduced prosocial behavior. Funnel and Doi plot asymmetry, however, was seen within the general video game exposure subgroup for aggressive behavior, but paradoxically, there was a preponderance of studies with smaller effects, which suggests that the true effect may have been underestimated by our analysis. For the other outcomes, we saw positive major asymmetry that could suggest (among other reasons) publication bias toward selective publications that reported a harmful effect of video game exposure. Therefore, even with the lack of significant associations with these outcomes, the presence of positive asymmetry (Figs. 1–5) strongly suggests that the point estimates themselves are spuriously positive.

Misconceptions About Meta-Analysis

Partial correlations

In their commentaries, Rothstein and Bushman (2015), Valkenburg (2015), and Boxer et al. (2015) argued that Ferguson's conclusions were "incorrect" because he employed adjusted effect sizes (partial correlations) instead of unadjusted/raw data (zero-order bivariate correlations) in his meta-analysis. We disagree with these comments for the following reasons:

1. The underlying concept of regression models is to adjust for confounding variables and provide a more accurate effect size estimate (unless the regression model was adjusted for the incorrect covariates/colliders). Therefore, if correctly adjusted partial correlations are pooled in a meta-analysis, the pooled estimate would be closer to the true underlying effect size in comparison with the pooled estimate that employed unadjusted effect sizes.
2. The main criticism against pooling adjusted correlations in meta-analyses is that each adjusted correlation may arise from a different regression model (adjusted for different covariates). The argument is that different sets of predictors lead to different interpretations of the partial correlation for each model. In the case of the meta-analysis presented here, we should not focus on the specific predictors, but rather on the larger constructs that the variables represent, which are all likely to be similar (Aloe, 2014, 2015; Aloe & Thompson, 2013). Examples of the use of adjusted effect sizes in behavioral science include the meta-analyses by Pratt and Cullen (2000) and Savage and Yancey (2008).
3. The variance of the partial correlation is larger than that of the bivariate correlation because the additional covariates lead to p fewer degrees of freedom, where p is the number of covariates adjusted for (Aloe & Thompson, 2013). By treating them as bivariate correlations, we assume a smaller variance:

$$\mathrm{var}(r_p) = \frac{(1 - r_p^2)^2}{n - p - 1}.$$

The variance structure of the partial correlation has the same form as that of the bivariate correlation, and thus Fisher's Z transformation can be used and the only difference is the incorporation of p in the denominator of the partial correlation, and when $p = 0$ then the two variances are identical (Aloe & Thompson, 2013). A simple Monte Carlo simulation over 1,000 iterations setting p between 1 and 10 from a uniform distribution confirmed that there was no appreciable change in the pooled effect size for aggression in comparison with leaving p out.

In summary, when synthesizing partial effect sizes, it seems useful to distinguish between similarity of models (similar set of predictors) and statistical homogeneity of effects (Aloe, 2015). In this context, similarity of models refers to the idea that the synthesized partial effect sizes arise from models that include a conceptually similar set of predictors (i.e., covariates as in the Angry Birds meta-analysis) and thus reflect the same population parameters, whereas the alternative would be to ensure statistical homogeneity of the effect between studies before pooling (Aloe, 2015). Therefore, there should not be a

Table 1. Meta-Analytic Results by Type of Video Game Exposure (Overall, General, and Only Violent Games) Using the Inverse Variance Heterogeneity and the Random Effects Model

Outcome measured	Number of studies	Pooled effect size estimate (IVhet model)	Pooled effect size estimate (RE model)	I^2 (%)	tau^2	Publication bias (LFK index)
Aggressive behavior						
Overall video game exposure	35	**0.059 (0.028 to 0.090)**	**0.049 (0.026 to 0.072)**	73	0.002	0.04 (No asymmetry)
General video game exposure	8	**0.061 (0.012 to 0.111)**	**0.045 (0.002 to 0.088)**	88	0.003	–2.38 (Major asymmetry)
Violent video game exposure	27	**0.055 (0.019 to 0.091)**	**0.051 (0.022 to 0.079)**	58	0.002	0.68 (No asymmetry)
Reduced prosocial behavior						
Overall video game exposure	12	0.068 (–0.019 to 0.155)	**0.063 (0.011 to 0.115)**[a]	52	0.004	–0.43 (No asymmetry)
General video game exposure	2	**0.072 (0.045 to 0.100)**	**0.072 (0.045 to 0.100)**	0	0.000	[b]
Violent video game exposure	10	0.058 (–0.022 to 0.138)	0.064 (–0.011 to 0.140)	60	0.009	–0.27 (No asymmetry)
Reduced academic performance						
Overall video game exposure	12	–0.004 (–0.046 to 0.038)	0.013 (–0.024 to 0.051)	54	0.002	3.88 (Major asymmetry)
General video game exposure	8	0.005 (–0.051 to 0.061)	0.025 (–0.023 to 0.073)	51	0.002	4.26 (Major asymmetry)
Violent video game exposure	4	–0.020 (–0.094 to 0.054)	0.000 (–0.070 to 0.070)	64	0.003	[b]
Depressive symptoms						
Overall video game exposure	16	0.018 (–0.039 to 0.076)	0.034 (–0.003 to 0.070)	80	0.003	2.63 (Major asymmetry)
General video game exposure	11	0.019 (–0.057 to 0.095)	0.039 (–0.011 to 0.089)	86	0.005	2.62 (Major asymmetry)
Violent video game exposure	5	0.013 (–0.022 to 0.049)	0.013 (–0.022 to 0.049)	0	0.000	2.47 (Major asymmetry)
Attention deficit symptoms						
Overall video game exposure	6	0.026 (–0.027 to 0.078)	0.034 (–0.018 to 0.085)	12	0.001	4.73 (Major asymmetry)
General video game exposure	3	0.096 (–0.038 to 0.230)	0.096 (–0.038 to 0.230)	0	0.000	[b]
Violent video game exposure	3	0.018 (–0.061 to 0.097)	0.034 (–0.043 to 0.110)	49	0.002	[b]

Note: Statistical significant pooled effect size estimates are in bold face. IVhet = inverse variance heterogeneity; RE = random effects.
[a]False positive association.
[b]LFK index reported for meta-analyses with at least 5 studies.

problem to synthesize partial effects representing similar models that are statistically heterogeneous as in this meta-analysis, even when we use the smaller variance that comes from ignoring the number of covariates.

The distribution of true effects

Rothstein and Bushman (2015) point out in their commentary that "In a random-effects meta-analysis, the mean overall effect is the average of a family of true effect sizes; the distribution of effects around this mean. . ." (p. 678), and we believe that this is a completely misleading statement. Rothstein and Bushman ignore the distinction between the model Ferguson chose for analysis of this dataset, and the mechanism by which the data came into being. An RE can be present in either of these roles, but the two roles are quite distinct. There is no reason to think the meta-analytical model and data-generation mechanism (model) are similar in form, but in their commentary, Rothstein and Bushman seem to suggest that the data-generation mechanism (model) is identical to the meta-analytical model Ferguson chose. Just because the analysis model includes an RE, it does not mean that the studies in a meta-analysis will necessarily arise as draws from some

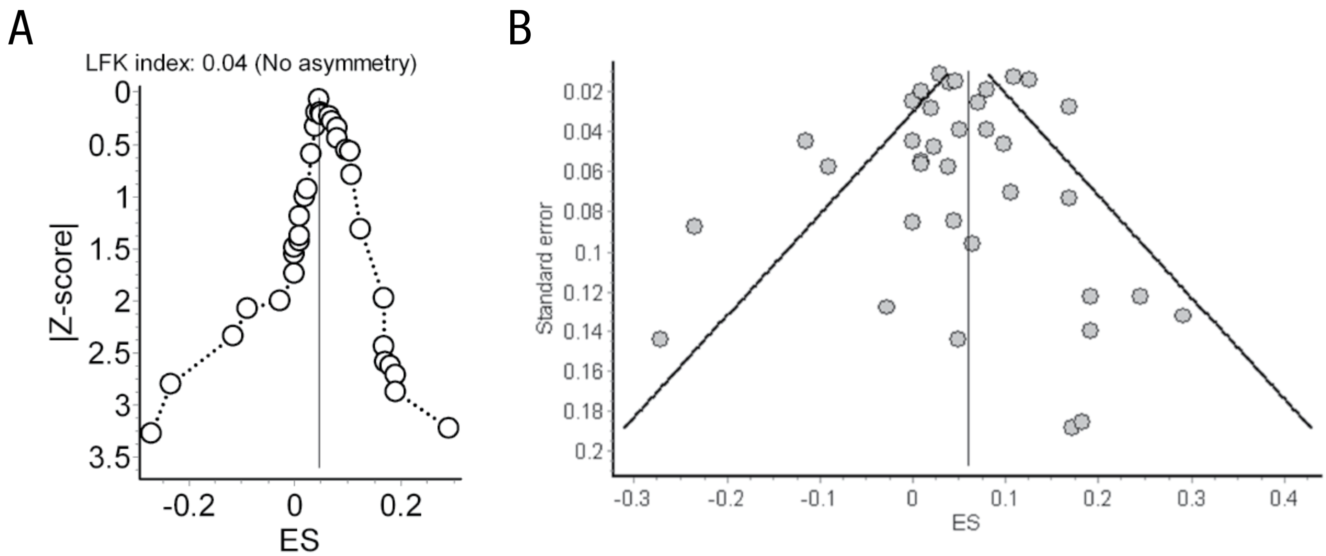

Fig. 1. Doi plot (A) and funnel plot (B) for aggressive behavior. ES = effect size.

mysterious RE distribution out there in the world. In fact, heterogeneity between studies arises from differences in patient populations, concomitant care, and measurement biases. We therefore have to agree with Hodges who states that this form of RE (alluded to by Rothstein and Bushman) is a hazardous distraction and, as hypothesized mechanisms for producing the data, the RE model for meta-analysis is silly; it is more appropriate to think of this model as a superficial description and something we choose as an analytical tool (Hodges, 2013)—but this choice for meta-analysis does not work because the study effects are a fixed feature of the respective meta-analyses and the probability distribution is only a descriptive tool. Given the aforementioned problem, the RE model results can suffer from several errors (depending on the mix of studies and heterogeneity, which is not that large in this meta-analysis) such as underestimation of the statistical error, spuriously overconfident estimates, and tendency toward exaggerating publication bias due to oversmoothing of the data to the point where weights are more or less equal, thus defaulting to an unweighted or arithmetic mean (Noma, 2011; Poole & Greenland, 1999). The scientific

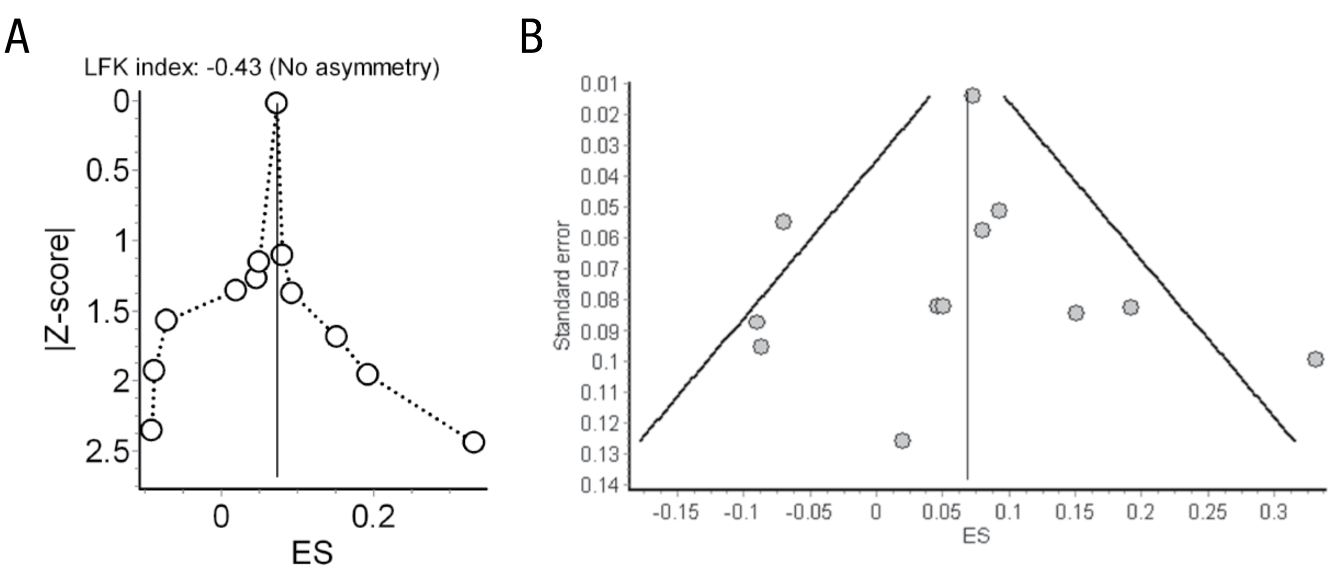

Fig. 2. Doi plot (A) and funnel plot (B) for reduced prosocial behavior. ES = effect size.

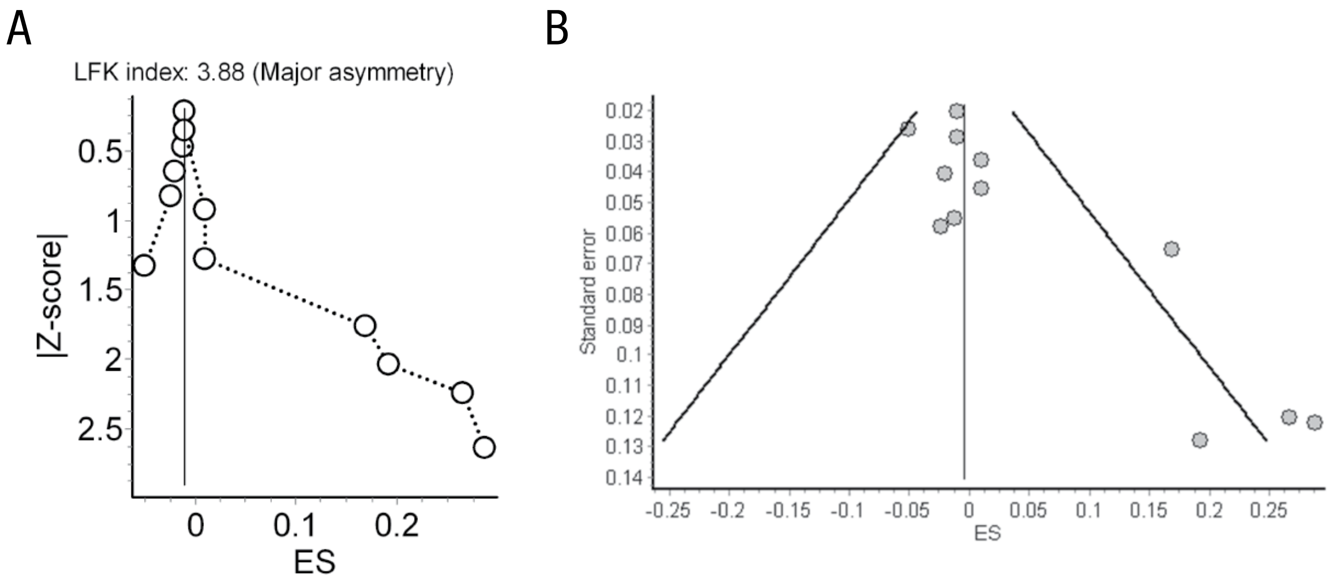

Fig. 3. Doi plot (A) and funnel plot (B) for reduced academic performance. ES = effect size.

community's fixation on the fallacy of the normal distribution of RE not only holds back scientific progress but also produces spurious results. We therefore used the IVhet model (Doi et al., 2015) for all the analyses presented here, though we also report RE results for comparison purposes and one (presumably) false positive association immediately pops up (overall videogame exposure and reduced prosocial behavior; Table 1).

Next, Rothstein and Bushman argue that Ferguson not reporting the prediction interval is a major problem. The 95% prediction interval has been suggested as a way to mitigate the known underestimation of the statistical error with the RE confidence interval and is defined as the expected effect of a treatment when it is applied within an individual setting and provides its bounds in 95% of the individual study settings. The prediction interval cannot replace the confidence interval, and it is based on the specific conceptualization of the underlying model as the RE summary treatment effect and utilizes the between-trial variance (Riley, Higgins, & Deeks, 2011). An assumption

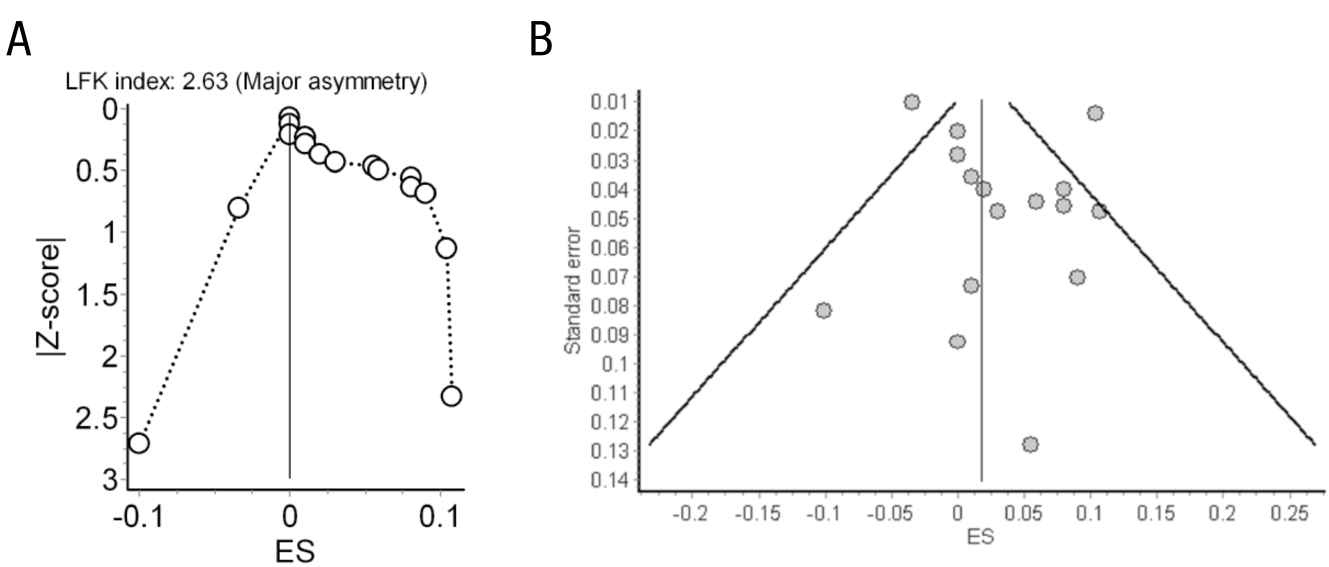

Fig. 4. Doi plot (A) and funnel plot (B) for depressive symptoms. ES = effect size.

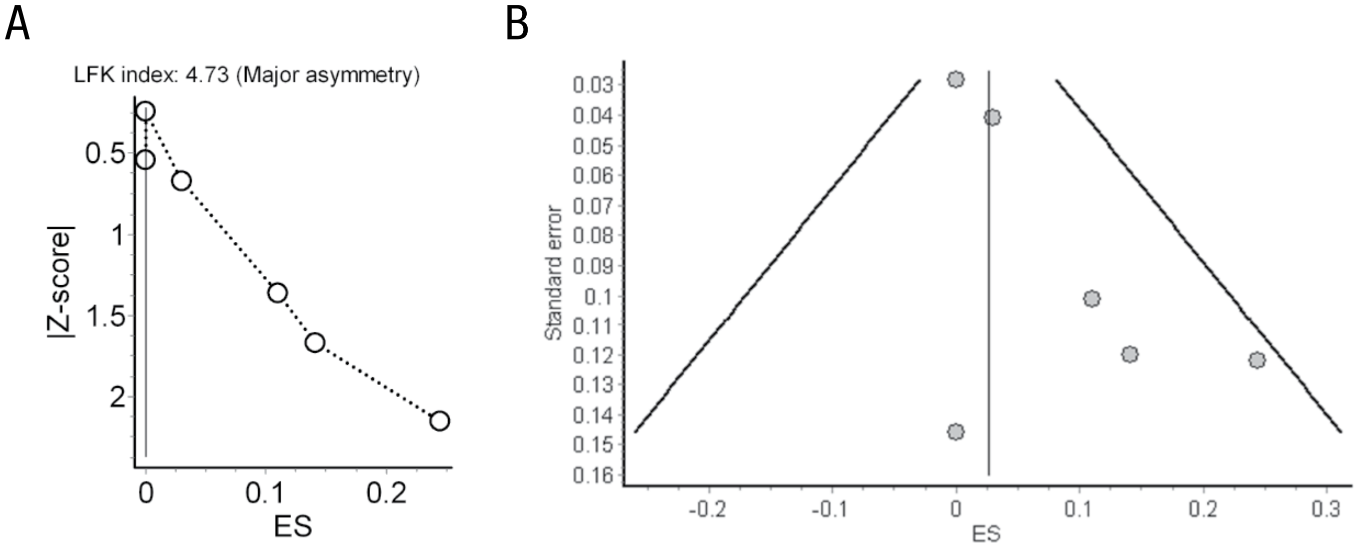

Fig. 5. Doi plot (A) and funnel plot (B) for attention deficit. ES = effect size.

behind this interval is that trials are considered more or less homogeneous entities and included patient populations and comparator treatments should be considered exchangeable. We therefore agree with Kriston who suggests that "if this is not the case (as with this meta-analysis), then prediction intervals are probably just as useless as random effect estimates" (Kriston, 2013, p. 4).

Do not get angry, instead give solutions

We agree with the researchers about some of the drawbacks in Ferguson's meta-analysis, but by no means do we think these disqualify Ferguson's findings. First, the use of a single coder should be addressed in subsequent studies, and at least two researchers should select the studies for inclusion and extract the data independently. Second, publication bias is a major problem for interpretation of meta-analyses, but we find it less plausible that the inclusion of imputed data can generate "bias-corrected estimates" based on the assumption that funnel plots should be symmetrical as it suggests that publication bias is the only source of such asymmetry. Instead, we provide an alternative graphical and a statistical method for detection of asymmetry rather than publication bias per se. Finally, if there is disagreement with the methods used by a researcher or if we have doubts regarding the methodology or results presented in a paper, the best option is probably to include a reanalysis of the data rather than a parallel commentary of criticisms, especially when (in our opinion) they are conceptually wrong.

In Conclusion

As suggested by Markey (2015), we reanalyzed the data without a preconceived opinion about the effects of video games on children mental health. We found a very small effect size for the association of video games and aggressive behavior that is similar in terms of reduced prosocial behavior both in the absence of funnel or Doi plot asymmetry. Paradoxically, even in the presence of asymmetry, which suggests publication bias favoring a harmful effect, our reanalysis found no association for the other three outcomes (reduced academic performance, depressive symptoms and attention deficit symptoms). We are quite confident that our results accurately demonstrate that violent and general video games exposure has a statistically significant, yet very small, effect size associated with aggression in children that in effect could represent no association. However, an infrequent effect modifier cannot be ruled out as a cause for the tiny positive correlations seen. If this is true, then conditioning on such a variable would cause the correlation to strengthen and a search for such variables should be undertaken in future research.

Acknowledgments

The authors are indebted to Christopher J. Ferguson for kindly providing us with the necessary data for this reanalysis.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

L. Furuya-Kanamori is funded by an Endeavour Postgraduate Scholarship (#3781_2014), an Australian National University Higher Degree Scholarship, and a Fondo para la Innovación, Ciencia y Tecnología Scholarship (#095-FINCyT-BDE-2014).

References

Aloe, A. M. (2014). An empirical investigation of partial effect sizes in meta-analysis of correlational data. *The Journal of General Psychology*, *141*, 47–64.

Aloe, A. M. (2015). Inaccuracy of regression results in replacing bivariate correlations. *Research Synthesis Methods*, *6*, 21–27.

Aloe, A. M., & Thompson, C. G. (2013). The synthesis of partial effect sizes. *Journal of the Society for Social Work & Research*, *4*, 390–405.

Barendregt, J. J., & Doi, S. A. (2015). *MetaXL User Guide version 3.1. p36-37.* Retrieved from http://www.epigear.com/index_files/MetaXL%20User%20Guide.pdf

Boxer, P., Groves, C. L., & Docherty, M. (2015). Video games do indeed influence children and adolescents' aggression, prosocial behavior, and academic performance: A clearer reading of Ferguson (2015). *Perspectives on Psychological Science*, *10*, 671–673.

Dersimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, *7*, 177–188.

Doi, S. A., Barendregt, J. J., Khan, S., Thalib, L., & Williams, G. M. (2015). Advances in the meta-analysis of heterogeneous clinical trials. I: The inverse variance heterogeneity model. *Contemporary Clinical Trials*, *45*, 130–138.

Ferguson, C. J. (2015a). Do angry birds make for angry children? A meta-analysis of video game influences on children's and adolescents' aggression, mental health, prosocial behavior, and academic performance. *Perspectives on Psychological Science*, *10*, 646–666.

Ferguson, C. J. (2015b). Pay no attention to that data behind the curtain: On angry birds, happy children, scholarly squabbles, publication bias, and why betas rule metas. *Perspectives on Psychological Science*, *10*, 683–691.

Gentile, D. A. (2015). What is a good skeptic to do? The case for skepticism in the media violence discussion. *Perspectives on Psychological Science*, *10*, 674–676.

Hodges, J. S. (2013). Random effects old and new. In J. S. Hodges (Ed.), *Richly parameterized linear models: Additive, time series, and spatial models using random effects* (pp. 285–302). Beaverton, OR: CRC Press.

Kriston, L. (2013). Dealing with clinical heterogeneity in meta-analysis. Assumptions, methods, interpretation. *International Journal of Methods in Psychiatric Research*, *22*, 1–15.

Markey, P. M. (2015). Finding the middle ground in violent video game research: Lessons from Ferguson (2015). *Perspectives on Psychological Science*, *10*, 667–670.

Noma, H. (2011). Confidence intervals for a random-effects meta-analysis based on Bartlett-type corrections. *Statistics in Medicine*, *30*, 3304–3312.

Poole, C., & Greenland, S. (1999). Random-effects meta-analyses are not always conservative. *American Journal of Epidemiology*, *150*, 469–475.

Pratt, T. C., & Cullen, F. T. (2000). The empirical status of Gottfredson and Hirschi's general theory of crime: A meta-analysis. *Criminology*, *38*, 931–964.

Riley, R. D., Higgins, J. P., & Deeks, J. J. (2011). Interpretation of random effects meta-analyses. *British Medical Journal*, *342*, d549.

Rothstein, H. R., & Bushman, B. J. (2015). Methodological and reporting errors in meta-analytic reviews make other meta-analysts angry: A commentary on Ferguson (2015). *Perspectives on Psychological Science*, *10*, 677–679.

Savage, J., & Yancey, C. (2008). The effects of media violence exposure on criminal aggression: A meta-analysis. *Criminal Justice and Behavior*, *35*, 772–791.

Valkenburg, P. M. (2015). The limited informativeness of meta-analyses of media effects. *Perspectives on Psychological Science*, *10*, 680–682.