Developmental Psychology

Is the Association Between Children's Baby Video Viewing and Poor Language Development Robust? A Reanalysis of Zimmerman, Christakis, and Meltzoff (2007)

Christopher J. Ferguson and M. Brent Donnellan

Online First Publication, July 15, 2013. doi: 10.1037/a0033628

Developmental Psychology

0012-1649/13/$12.00 DOI: 10.1037/a0033628

BRIEF REPORT

Is the Association Between Children's Baby Video Viewing and Poor Language Development Robust? A Reanalysis of Zimmerman, Christakis, and Meltzoff (2007)

Christopher J. Ferguson
Stetson University

M. Brent Donnellan
Michigan State University

Zimmerman, Christakis, and Meltzoff (2007) reported that exposure to *Baby Einstein* videos was negatively associated with language development. The current study uses the Zimmerman et al. (2007) data set to replicate and extend the original analyses. Caregivers of 392 children aged 6 to 16 months and 358 children aged 17 to 27 months reported on media exposure, language development, and control variables related to child/parent interaction and demographic characteristics. Results indicated that exposure to baby videos could be construed as positive, neutral, or negative depending upon the statistical analysis. The effect size estimates were generally negligible across analyses. Exposure to educational programming tended to be positively related to language development. Infants exposed to no media actually had lower levels of language development compared to infants with some exposure. Given these results, the Zimmerman et al. (2007) data set does not permit strong inferences about a connection between exposure to media and language development in young children. These results also highlight recent concerns over methodological flexibility and the possibility of increased Type I errors (false positives) in the psychological literature.

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

The American Academy of Pediatrics (AAP; 2011) discourages the exposure of children younger than 2 years of age to electronic media such as television programs and DVDs. The AAP is concerned that media exposure may contribute to language delays and potential attention problems in young children. Although the AAP acknowledged the need for more research, it cited a 2007 correlational study by Zimmerman, Christakis, and Meltzoff (2007) as supporting its position. The Zimmerman et al. (2007) study evaluated associations between different types of media exposure and language development. The authors found no evidence for media effects for children between 17 and 24 months; however, they found an association between exposure to baby videos (i.e., videos and DVDs, such as *Baby Einstein*, marketed to parents of infants) and diminished language development for children 8 to 16 months old. The University of Washington (2007) issued a press release accompanying this study and quoted Christakis as claiming that baby videos "may in fact be harmful" (para. 9). The public reaction to this study may even have prompted the Walt Disney Company to offer refunds to parents who had purchased the popular *Baby Einstein* series of video products in 2009 (Lewin, 2009).

All told, the impact of the Zimmerman et al. (2007) article seems to have extended beyond the research literature concerned with media influences on cognitive development. However, this article has proven controversial (Lewin, 2010), and other research fails to support a clear negative association between media exposure and cognitive outcomes (Foster & Watkins, 2010; Schmidt, Rich, Rifas-Shiman, Oken, & Taveras, 2009; Stevens & Mulsow, 2006). One recent review on this topic concluded that "it is premature either to condemn television and video material as a source of harm to the developing infant brain or to promote it as a viable source of early learning" (Courage & Setliff, 2009, p. 76). Thus, additional research is needed on this practically important topic. The goal of this study is to reanalyze the Zimmerman et al. (2007) data set to evaluate their conclusions in light of recent concerns about methodological practices in psychological research.

Indeed, Foster and Watkins (2010) noted that reanalysis is especially valuable when research findings are used to inform public health policies and parental beliefs and practices. In line with their argument about the value of reanalysis, Foster and Watkins reexamined the association between early exposure to television and later attention problems reported by Christakis, Zimmerman, DiGiuseppe, and McCarty (2004). Foster and Watkins argued that there was no "meaningful relation" (p. 368)

Christopher J. Ferguson, Department of Psychology, Stetson University; M. Brent Donnellan, Department of Psychology, Michigan State University.

Correspondence concerning this article should be addressed to Christopher J. Ferguson, Department of Psychology, Stetson University, 421 North Woodland Boulevard, DeLand, FL 32723. E-mail: CJFerguson1111@aol.com

between exposure and attention problems and suggested that the earlier findings were not robust when controlling for maternal academic achievement and early poverty status. Similarly, the current report evaluates whether the Zimmerman et al. (2007) conclusions are robust across different ways of approaching the analyses.

False Positives and Methodological Flexibility in Research

Fanelli (2010) found that almost all papers (91.5%) in the behavioral sciences publish results that confirm the a priori hypotheses of the authors. This trend toward confirmation appears to be increasing over time (Fanelli, 2012). One explanation for the preponderance of positive findings in the literature is that hypothesis-supportive results are published more frequently than null results—a phenomenon known as publication bias (Ferguson & Brannick, 2012). This trend toward confirmation raises concerns in light of a growing sense of unease over the prevalence of false positive results in the literature (e.g., LeBel & Peters, 2011). Ioannidis (2005) even proclaimed that most published findings are false.

One factor that seems to contribute to false positives has been termed the problem of "methodological flexibility" or "researcher degrees of freedom" (LeBel & Peters, 2011; Simmons, Nelson, & Simonsohn, 2011). Methodological flexibility originates from the large number of choices that are inherent when conducting, analyzing, and reporting scientific studies. The concern is that researchers may approach these decisions with particular biases, such as a preference for statistically significant results or a preference for results that confirm initial expectations. The idea of privileging confirmatory results over ambiguous or disconfirmatory results challenges the dominant perception of psychological research as following a falsificationist paradigm. We should note that any preference for confirmatory results is not necessarily conscious or inherently malicious, but it seems to interfere with the overarching goal of getting things right when conducting research (see, e.g., Nosek, Spies, & Motyl, 2012).

In light of worries about the integrity of the psychological literature, there is an increasing amount of attention to practices that can be implemented to reduce the prevalence of false positives or Type I errors. One recommendation involves making raw data available for reanalysis so that other researchers (presumably those with different expectations) can evaluate the robustness of focal effects to different model specifications and alternative statistical analysis (e.g., Kashy, Donnellan, Ackerman, & Russell, 2009; Simonsohn, 2012; Wicherts & Bakker, 2012). This approach provides an opportunity for the psychological community to subject findings to additional critical evaluation.

As it stands, there are cases in which studies were reanalyzed by outside groups leading to different results and conclusions. The previously discussed Foster and Watkins (2010) article is a relevant example for the current study. Likewise, a longitudinal study suggesting that watching sexually charged media is related to teen sexual behavior (Brown et al., 2006) was reanalyzed by Steinberg and Monahan (2010). The reanalysis suggested that the focal association was essentially eliminated when other factors were controlled (see also Brown, 2011; Collins, Martino, & Elliott, 2011; and Steinberg & Monahan, 2011, for full discussion). Such examples are not limited to research concerning media influences on developmental outcomes. Personality data suggesting that the Big Five personality traits yielded a substantive general personality factor (Van der Linden, Scholte, Cillessen, te Nijhenhuis, & Segers, 2010) were also reanalyzed and challenged (de Vries, 2011; but see also van der Linden, 2011). We suspect that a pattern of analysis, reanalysis, and potential rejoinder is actually quite common in psychology.

Differences in results often boil down to differences in the specification of regression models. The basic approach is to include controls for "third variables" that may account for any connection between a focal predictor and outcome variables. This practice is useful because it helps to rule out alternative explanations for particular associations. The ideal situation occurs when the focal result is consistent across a number of different models with different covariates (i.e., different model specifications). In other words, the ideal situation occurs when the effect size for the focal predictor is statistically distinguishable from zero and largely unchanged across a wide variety of regression equations with different covariates (Keith, 2006). This approach is consistent with the data-driven approach advocated by Greenwald, Pratkanis, Leippe, and Baumgardner (1986).

A more worrisome situation occurs when interpretations regarding the focal predictor shift depending on the set of covariates included in the model. This increases the likelihood that a researcher may draw incorrect interpretations from a model (Zientek & Thompson, 2006). A particularly troubling scenario may occur when a covariate or a specific set of covariates is required to obtain the hypothesized effects. These situations are especially problematic when model specification is not guided by strong theoretical and conceptual reasoning, and there is little agreement over the appropriate suite of so-called "control" variables (see Azen, Budescu, & Reiser, 2001, for discussion of this concern). This means there are fewer compelling reasons for having a preference for a certain set of results over another besides initial biases. To potentially address these kinds of concerns, Simmons et al. (2011) recommended that authors report zero-order effects as well as the conditional effects from multivariate analyses. They also suggested that authors provide a reasoned and theoretically grounded approach to model specification. This level of transparency and explicitness facilitates criticism of the reported analyses.

The Current Reanalysis

As it stands, we believe there are compelling reasons to revisit the Zimmerman et al. (2007) article because of concerns over researcher degrees of freedom. The most basic reason is that the zero-order effect for baby video exposure was not statistically significant at the conventional alpha level of .05. Although the correlations were not reported in the original report, they were provided in response to an e-mail request (F. Zimmerman, personal communication, August 25, 2007). The provided correlation was –.07 for baby video exposure and language development in the younger cohort (95% CI [–.17, .03]; $N = 384$; $p = .17$, two-tailed), whereas the coefficient was .04 for the older cohort (95% CI [–.06, .15]; $N = 345$; $p = .42$). These two coefficients were not significantly different from each other ($z = 1.53$; 95% CI for the difference: [–.25, .03]). In short, the Zimmerman et al. report is an example of an analysis in which the inclusion of covariates is needed to generate a statistically significant result for exposure to baby videos.

The main analyses in the Zimmerman et al. (2007) article included a number of covariates (approximately 20; see their Table 2, p. 367). It is likely that space requirements for the journal precluded a lengthy rationale for each variable. Some covariates such as parental income and parental education are straightforward, given that socioeconomic status might be related to both exposure to baby media and language development. However, the justification for including other covariates is not as straightforward or clear-cut. For example, Zimmerman et al. (2007) used dependent language variables that were transformed to percentile norms using different standards for girls and boys. Although this was not completely unreasonable, Zimmerman et al. (2007) then included sex as a covariate, which makes the interpretation of this variable and other variables challenging, given that the dependent variable (DV) should have already been adjusted for sex. The original authors also included a set of covariates to address exposure to different types of media (i.e., educational programming, noneducational programming, and adult TV), along with a separate variable to capture no media exposure. This set of presumably correlated media exposure variables might produce artifacts and ultimately make it difficult to interpret any particular media exposure variable in isolation. The central issue is that the parameter estimates for any single predictor are conditional on the other variables in the regression model.

In sum, we believe there is clear value in reexamining the Zimmerman et al. (2007) data set using the same general linear regression approach used by those authors to test the sensitivity of the baby video results to different model specifications. In addition, we provide an explicit rationale for the inclusion of control variables based upon previous research and findings in the media exposure literature. We also wanted to provide effect size estimates in a metric that is easily understandable (Ferguson, 2009) so that the magnitude of the findings could be more clearly grasped by psychological researchers.

Method

Procurement of the Data

The original data file was requested from the University of Washington (UW) under a public records request. The file in SPSS format was delivered to us as requested. This consisted of the raw data for 1,009 cases (although Zimmerman et al., 2007, listed the complete data set as including 1,008 cases). We constructed composite variables from this file. Analyses were restricted to the 750 families who were administered measures of language development and media exposure. We also received a copy of the file from William Clark (cocreator of the *Baby Einstein* educational videos for infants), who received the copy pursuant to a lawsuit against UW. The two files were virtually identical, except that the Clark version included calculated ages and omitted dates of birth to avoid identifying data. Precise sample sizes fluctuated across analyses because of missing data; these are noted where appropriate.

Participants and Procedures

Participants were the caregivers (87.9% mothers) of 392 children aged 6 to 16 months and 358 children aged 17 to 27 months; caregivers responded to a phone interview assessing media exposure, language development, and other variables related to parent–child interaction and demographic characteristics (see Zimmerman et al., 2007, for information about participant characteristics and recruitment procedures). There were several cases of children whose ages were slightly outside the age ranges reported in the Zimmerman et al. (2007) article (e.g., there were several children aged 25–27 months in the file). However, including these cases did not alter results. The results reported in this article therefore include all children in the data file.

This sample was fairly homogenous when considering the reports of the racial/ethnic group membership of the child. Of the sample with language data, 647 (86.3%) identified their child as Caucasian, nine (1.2%) as African American, six (0.8%) as Native American, 20 (2.7%) as Hispanic, seven (0.9%) as Asian/Pacific Islander, 58 (7.7%) as multiethnic, and three (0.4%) as other or refused to answer. In a separate ethnicity variable, 704 (93.9%) reported their child was non-Hispanic and 41 (5.5%) reported their child was Hispanic. Thus, there was some disagreement between the two variables concerning the Hispanic designation. If it can be assumed that any individual who endorsed being Hispanic on either variable was Hispanic, then the number of Hispanic children was 42 (5.6%). Thus, most (but not quite all) caregivers who identified their offspring as Hispanic in terms of race also identified ethnically as Hispanic, but not the inverse.

Measures and Variables

We describe the measures and variables here briefly, as these are described in more detail in the original report (Zimmerman et al., 2007). Except where noted below, we followed the procedures described in Zimmerman et al. regarding the construction of study variables.

Media exposure. Four media exposure variables were calculated following Zimmerman et al. (2007, p. 365). Caregivers were asked how much time on a typical weekday and on a typical weekend day that children were exposed to different types of media. Participants responded by giving the number of hours and minutes. We transformed these reports into a composite exposure per week variable by multiplying the typical weekday estimate by 5 and adding that product to twice the typical weekend estimate. This weekly total was divided by 7 to provide an estimate of daily exposure in minutes. Media exposure questions covered baby DVDs or movies, adult TV (e.g., the nightly news), children's educational TV (e.g., *Sesame Street*), children's educational DVDs or videos (e.g., *Sesame Street*), children's noneducational TV (e.g., *SpongeBob SquarePants*), and children's DVDs or videos of movies (e.g., *The Lion King*); however, Zimmerman et al. (2007) collapsed the last four questions into educational and noneducational media exposure. Thus, the four final media variables were baby DVDs/movies, adult TV, children's educational shows, and children's noneducational shows. We constructed a separate no-media variable (1 = no exposure; 0 = nonzero exposure) which was analyzed independently of the other four categories (262 respondents reported that their child did not have any media exposure based on responses to the variables considered in this study). We also created a summary media exposure variable that was simply the total time of media exposure.

All four media exposure variables had noticeably nonnormal distributions, and there were a number of potential outliers for each variable. The baby video exposure variable is a good illustration. It was skewed (skew: 5.58, SE = 0.09) and the median score for exposure to baby videos was zero, whereas the average score was 9.46 min per day (SD = 23.14; 5% trimmed M = 5.81; 509 cases reported zero exposure). There were 10 cases with z scores above 3. The two most extreme cases reported exposure levels of 257.14 min per day on average. It is possible that these cases may have been the result of careless responding or data recording errors (e.g., the most extreme responses indicate that children were exposed to baby videos for over 4 hr per day), but we have no way of verifying this possibility. Such variables present challenges for traditional data analyses. Although ordinary least squares (OLS) regressions can handle variables with nonnormal distributions in principle, outliers and nonnormal distributions often create challenges in terms of leverage and distorted standard errors (Wilcox, 2012; see also Stevens, 1984).

To address concerns about the distributions of the media exposure variables, we conducted a square-root transformation in an attempt to make the distributions more normal. This same strategy was followed by Anderson, Huston, Schmitt, Linebarger, and Wright (2001) when they investigated associations between early childhood exposure to television programming and adolescent outcomes (p. 23). This transformation is often used for count variables, but transformations do not always reduce the impact of outliers (Wilcox, 2012). We describe other attempts to address outliers in the Results section.

Language development. Language development was assessed using the Communicative Development Inventory (CDI; Fenson et al., 2000). Zimmerman et al. (2007) made separate inquires about receptive and expressive language for children in the 6- to 16-month-old range (αs = .97 and .96, respectively), whereas there was a single language score for children in the 17- to 27-month-old range (α = .98). Considering the 6- to 16-month-old battery, the receptive language variable was based on an affirmative caregiver response to whether the child understood a word, whereas the expressive language variable was based on whether the child used the word. Caregivers of the younger cohort were presented with 89 words. Caregivers of older children were asked about receptive language for 100 words.

The correlation between the receptive and expressive language variables was .61. Zimmerman et al. (2007) analyzed only receptive language scores for children in the 6- to 16-month-old range. It is not clear why results for expressive language were not reported, although the skew and restricted range for expressive language in young children is a possible explanation. As it stands, however, there was a fairly considerable range in responses (range: 0–74 words, M = 8.48, SD = 11.34, Mdn = 4.00, skewness = 2.32, kurtosis = 6.57) but zero words was a common response (15.4%). On balance we believed this variable was worth considering, so it was included in the current report. Using a square-root-transformed version of the expressive language variable to reduce skew did not substantially change our results regarding media influences. Data for the receptive language variable for the younger sample was more normally distributed (range: 0–87, M = 34.37, SD = 20.74, Mdn = 34, skewness = .15, kurtosis = –.965), with few zero-words responses (1.3%). The language variable for the older toddlers was likewise more normally distributed (range: 0–96, M = 36.86, SD = 33.00, Mdn = 33, skewness = .382, kurtosis = –.934) with few zero-words responses (1.1%). Zimmerman et al. (2007) used age and gender-normed percentile scores; however, percentile scores do not have optimal psychometric properties (Glass & Hopkins, 1996). One of the major limitations of percentile scores is that a presumably interval-level variable is rescaled to an ordinal ranking according to a reference distribution based on age and gender. Percentiles are often useful for providing descriptive results to consumers of tests, but the use of percentile scores in standard statistical tests may result in spurious results and is thus not recommended (Brown, 1983; Glass & Hopkins, 1996; Mehrens & Lehmann, 1991). The use of raw scores in a regression equation results in a straightforward interpretation, given that the metric of the dependent variable is the number of words. Unstandardized regression results are available upon request.

Covariates. Zimmerman et al. (2007) incorporated a number of covariates in their analyses including child age, gender, race/ethnicity, household income, maternal education, and paternal education. Zimmerman et al. (2007) also included parental interaction variables related to music, storytelling, and reading. These latter variables were dichotomized in the original article, but we left them in original form to maximize variance (see MacCallum, Zhang, Preacher, & Rucker, 2002). Zimmerman et al. acknowledged that the paternal education variable was missing for a significant number of participants. They stated that they replaced these missing values with the modal response. This approach is not recommended for handling missing data, and we did not use this variable in our analysis. As noted earlier, we encountered problems with the race/ethnicity variable; we explain below how this was handled.

Data analysis plan. We proceeded to conduct analyses in a sequential fashion to determine whether conclusions from the Zimmerman et al. (2007) data set change when different covariates are included in the model. We began by considering bivariate correlations between media viewing and the language development scores of children. Next we considered OLS regressions using a reduced set of covariates involving age and gender of the child in addition to the four media exposure variables when predicting children's language development. This was undertaken given our concerns about the impact the correlated media exposure variables might exert on the analyses. We then approximated Zimmerman et al.'s original OLS regression, including the covariates described above. We submit that a robust negative association between exposure to baby media and language development across these analyses would strengthen the arguments of Zimmerman et al. (2007), whereas inconsistencies would weaken the persuasiveness of their conclusions.

Differences Between Zimmerman et al. (2007) and the Current Analysis

In this section, we summarize the differences between our analysis and that of Zimmerman et al. (2007). First, we used raw language variables, whereas they used normed percentile scores. This was based on concerns about percentile scores and the fact that the sample used for norming the CDI short form was not collected using probabilistic-sampling methods (see Fenson et al., 2000), so it is unclear whether the published norms actually have validity. Second, because the media exposure variables were nonnormal with a considerable number of zeros, we used a square root transformation and took steps to deal with outliers. These kinds of considerations apparently were not taken into account by Zimmerman et al. (2007), but we also considered results using the raw variables. Third, we examined whether those children exposed to no media actually score higher on language development, to test whether these data are consistent with the overall AAP recommendations about media exposure.

Last, we provide an explicit theoretical rationale for the inclusion of covariates including age, gender, parental interaction, maternal education, and family income. Some control variables used by Zimmerman et al. (2007) appeared to be problematic in terms of data availability (paternal education) or were not clearly based on a question in the survey instrument (premature birth).[1] Others seemed to lack a clear theoretical rationale (site of data collection). These kinds of variables were not included in our models. Age naturally relates to language proficiency in very young children and is therefore critically important to include, as is gender, given that females tend to learn language faster than males (Lovas, 2011). Parental interaction has been found to be associated with toddler literacy (Weigel, Martin, & Bennett, 2006), and thus variables related to parental interaction would seem to be well justified as inclusion as covariates. Schmidt et al. (2009) found that television viewing is confounded with maternal education and household income, and these variables are likely to relate to language development, so these also seem like important controls (see also Anderson et al., 2001). We sought to include covariates in line with previous research and those with straightforward concerns regarding confounding. Covariates that lacked precedence or theoretical rationale for inclusion were excluded.

Results

Table 1 presents the bivariate correlations between media exposure variables and the language development variables. Positive correlations were evident between all media exposure variables. Moreover, media exposure variables were positively correlated with language variables in the younger sample but not the older sample. Avoidance of media was negatively correlated with both language development scores for children ages 6 to 16 months. Using nontransformed media variables did not impact these conclusions. Baby video viewing was positively associated with expressive language at 6–16 months (r = .11, p = .029) but not receptive language at 6–16 months or language development from 17 to 26 months (rs = .03 and .06, respectively, $p > .05$).

These zero-order correlations should be interpreted with caution. For instance, older children may be exposed to more media and have better language development than younger children. To be sure, age in months was correlated with total time spent viewing media (r = .31, $p < .001$), and age was associated with receptive and expressive language at 6–16 months and total language at 17–27 months (r = .60, .47, .44, respectively, all ps < .01). Gender of the child (1 = girl, 0 = boy) was unrelated to total media exposure (r = –.03) but was related to the three language development variables (r = .11, .12, and .19, respectively, all ps < .05). These analyses suggest that age is a critical control variable.

Controlling for Age and Gender

We used OLS regression results to evaluate the association between media exposure and language development, controlling for age and gender of the child. First, we considered separate regression models for each media exposure variable. In each case age and female gender were significant predictors of language development scores at all ages. These results are presented in Table 2.

Most media exposure variables had negligible or positive associations with language development scores when entered into separate regression equations. Exposure to educational and noneducational media was positively associated with language development, but only for younger children. No other media variables were significant predictors of language development. There was no evidence for an association for exposure to baby media and language at $p < .05$. Results did not differ at this stage using either transformed or nontransformed media variables. Collinearity diagnostics indicated a virtual absence of multicollinearity at this point, with variance inflation factors (VIF) around 1.00 and tolerances also around 1.00. In regression models using the binary no-media-exposure variable, the effect was negative for receptive language (β = –.13, p = .001) and expressive language (β = –.12, p = .007) for the younger sample. In other words, no exposure to any media was associated with lower scores for receptive and expressive language.

When the four transformed media exposure variables were entered together in the same OLS regression model (see Table 3), there were potential indications of multicollinearity for children ages 6–16 months (highest VIF = 2.20, lowest tolerance = .45). This approach also changed the results somewhat. Exposure to educational media retained its positive association with language development for younger children, whereas exposure to noneducational media did not. The effects for educational media exposure were generally at or below those suggested by some scholars for interpretability or practical significance (Ferguson, 2009). The effects for the baby DVDs/video variable were not statistically significant. However, when using nontransformed media variables, baby videos were significantly associated with receptive language development scores for younger children (β = –.09, p = .047, vs. β = –.05, $p > .05$, for the transformed variable), although the effect size for this association was arguably trivial. Further, this result may have depended on the multicollinearity between media variables. Removing exposure to children's educational media, for instance, reduced the baby-viewing video effect to nonsignificance (β = –.06, $p > .05$). The baby video effect was only statistically significant when children's educational videos were included in the model and only when using nontransformed (i.e., raw) media variables. Regressing receptive language on the raw baby video exposure variable and the sum of all other

[1] Although Zimmerman et al. (2007) claimed to control for premature birth, we could find no variable in the data set or question in the survey instrument pertaining to this variable. When the lead author was contacted about this, he stated he no longer remembered how this variable was handled (F. Zimmerman, personal communication, May 10, 2012), which is reasonable given the time interval. Regardless, it was not available to us for reanalysis.

Table 1
Bivariate Correlations Between Media Variables and Language Development

Variable	1	2	3	4	5	6	7	8
1. Receptive language, 6–16 mos.	1.00	.61**		.09	.10*	.29**	.19**	−.22**
2. Expressive language, 6–16 mos.		1.00		.14**	.11*	.33**	.24**	−.19**
3. Language 17–27 mos.			1.00	.03	.09	.09	.05	−.08
4. Baby DVDs/videos				1.00	.11*	.28**	.20**	−.44**
5. Adult TV					1.00	.42**	.37**	−.33**
6. Child educational						1.00	.54**	−.61**
7. Child noneducational							1.00	−.44**
8. No media exposure								1.00

Note. Media exposure variables (Variables 4–7) were square root transformed.
* $p < .05$. ** $p < .01$.

media exposure variables along with gender and age also yielded a nonsignificant baby video result (β = –.06, $p > .05$).

OLS Regressions With All Control Variables

We then conducted regression analyses similar to those reported by Zimmerman et al. (2007). Table 4 presents the standardized regression coefficients for the media exposure variables entered together with our control variables (age, gender, parental interaction variables, income, maternal education). Multicollinearity may have been a concern with children ages 6–16 months (highest VIF = 2.28, lowest tolerance = .44) but was less of a concern for older children (highest VIF = 1.35, lowest tolerance = .74). Results for this set of analyses were not notably different from the previous analyses that only controlled for age and gender (i.e., Table 2). As with the previous regression results, using transformed versus nontransformed media variables creates slight fluctuations in the regression weights. These small differences were enough to change the binary statistical significance decision between the two models. This fact can result in potentially different conclusions regarding the relevance of the baby video exposure variable when predicting receptive language for the younger group. As was the case in the previous section, removing children's educational video exposure reduces the effect for the nontransformed baby video variable to nonsignificance (β = –.08, p = .077). We also ran the OLS regressions with the no-media-exposure binary variable as the sole media predictor. In these analyses, the standardized regression coefficients for no media were statistically significant for receptive and expressive language development in younger children (βs = –.10 and –.09, respectively, both ps < .05) but not for language development in older children (β = –.03). A similar pattern occurred when using the square-root-transformed sum of media exposure variable.[2]

Table 2
Separate Regression Analyses for Each Media Exposure Category Considered Individually and Controlling for Age and Gender

Predictor variable	Receptive 6–16 mos.	Express 6–16 mos.	Language 17–27 mos.
Baby DVDs/video	.03/−.01	.09†/.08	.04/.04
Adult TV	.08*/.02	.10*/.03	.06/.06
Child educational	.18**/.16**	.25**/.27**	.05/.05
Child noneducational	.13**/.14**	.20**/.25**	.01/.01
No media exposure	−.13**	−.12*	−.03

Note. All values are standardized regression coefficients (i.e., betas). In this table, media exposure variables were entered separately, controlling only for age and gender. Numbers to the left of the backslash are for transformed media variables. Variables to the right are for nontransformed raw variables. Receptive = receptive language; Express = expressive language.
† $p = .05$. * $p < .05$. ** $p < .01$.

Ancillary Analyses

Outlier analyses. We were concerned that the raw variable analyses are sensitive to outliers; this can be seen in Figure 1, which shows the scatter plot of expressive language and baby video exposure for the younger cohort with a nonparametric regression line superimposed using locally estimated scatter plot smoothing (bandwidth set to .3). The most extreme value appears to be a substantial outlier and one that potentially exerts leverage on the analysis, given that this case also had an extremely low value on expressive language (i.e., 1.0).

To address potential outliers, we removed the 10 scores that were three standard deviations above the mean (i.e., values above 78.85) and repeated the raw variable analyses. This amounted to discarding four cases for the 6- to 16-month-old language variables and six cases for the 17- to 27-month-old language variable. Windsorizing rather than discarding these cases had no appreciable influence on our results. The previously significant negative association with receptive language was not statistically significant in the model with gender, age, and the four media exposure variables (β = –.04, p = .30) and in the model with all controls listed in Table 4 (β = –.06, p = .17). Thus, we found that a few outliers

[2] Given that specific covariates appeared capable of influencing the regression coefficients enough to produce differing decisions as per statistical significance, we also employed the model building approach recommended by Carta et al. (2001). Briefly, we found that of the social covariates (reading, storytelling, music, income, and maternal education), only reading and storytelling were significant covariates. The inclusion or noninclusion of the reading variable in the regression model influenced the statistical significance of baby videos when predicting receptive language. This is consistent with our conclusion that the statistical significance of the baby video exposure variable depends on the specific suite of covariates included in the model. Full details of these analyses are available on request.

Table 3
Regression Analyses With All Media Exposure Variables Entered Simultaneously and Controlling for Age and Gender

Predictor variable	Receptive 6–16 mos.	Express 6–16 mos.	Language 17–27 mos.
Baby DVDs/video	−.05/−.09*	−.01/−.04	.03/.03
Adult TV	−.03/−.04	−.06/−.06	.06/.06
Child educational	.19**/.15**	.23**/.21**	.03/.03
Child noneducational	.04/.07	.08/.14*	−.03/−.02

Note. All values are standardized regression coefficients (i.e., betas). In this table, media exposure variables were entered together, controlling for age and gender. Numbers to the left of the backslash are for transformed media variables. Variables to the right are for nontransformed raw variables. Receptive = receptive language; Express = expressive language.
* $p < .05$. ** $p < .01$.

seemed to have an influence on the conclusions drawn about the raw baby video exposure variable.

Controlling for race/ethnic group membership. The relatively small number of minority participants and the ambiguity concerning the reporting of Hispanic race/ethnicity suggests that a meaningful evaluation of race/ethnicity with this data set difficult. Nonetheless, we used several strategies to investigate the potential impact of ethnicity on our conclusions. We first reran the full model including a dummy-coded variable for Caucasian non-Hispanics versus ethnic minorities. These results are presented in Table 5. As can be seen, our results did not significantly change with minority status controlled. Again, a discrepancy between transformed and nontransformed media variables was evident for younger children (without outliers) but not for older children (without outliers). Then, we used separate dummy-coded variables for ethnic groups (Caucasian, African American, Native American, Hispanic, Asian, and multiethnic). To save space we do not report the specific effect sizes, but these variables had little impact on the media variables (e.g., β for baby videos was –.01, $p = .876$). We also reran analyses using the Caucasian sample only, and results were the same as those previously reported. Thus, ethnicity does not seem to influence the outcomes for media variables.

Table 4
OLS Regression With All Control Variables

Predictor variable	Receptive 6–16 mos.	Express 6–16 mos.	Language 17–27 mos.
Age	.56**/.57**	.41**/.43**	.42**/.42**
Gender (girl)	.10*/.10*	.11*/.10*	.21**/.20**
Reading	.11*/.11**	.06/.06	.08/.08
Stories	.12**/.12**	.04/.03	.18**/.18**
Music	.00/.01	−.03/−.02	.09/.08
Income	.01/.01	−.10*/.09	.06/.06
Maternal education	−.06/−.07	−.06/−.06	.08/.08
Baby DVDs/video	−.07/−.10*	−.03/−.05	.01/.04
Adult TV	−.04/−.04	−.07/−.06	.09/.10*
Child educational	.18**/.12*	.22**/.18**	.02/.02
Child noneducational	.04/.08	.07/.13*	.02/.02

Note. All values are standardized regression coefficients (i.e., betas). Numbers to the left of the backslash are for transformed media variables. Variables to the right are for nontransformed raw variables. Receptive = receptive language; Express = expressive language.
* $p < .05$. ** $p < .01$.

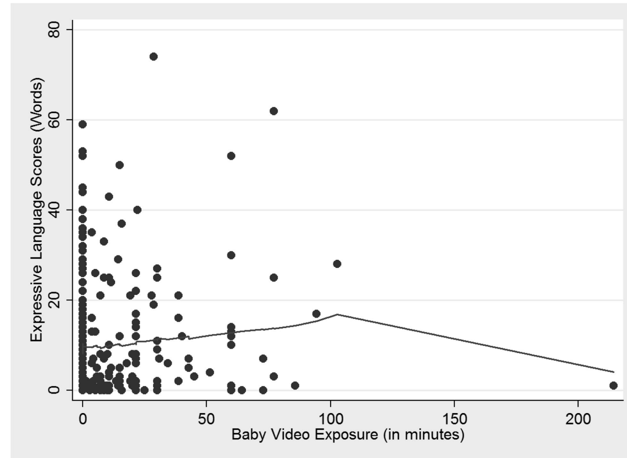

Figure 1. Scatter plot showing baby video exposure and expressive language scores. Line is smoothed with locally weighted least squares smoothing using Stata 12.0 with bandwidth set to .3.

Discussion

Zimmerman et al. (2007) concluded that there was evidence of a large negative association between exposure to baby media and vocabulary acquisition for children ages 8 to 16 months. The goal of the current study was to critically evaluate this conclusion. Our analyses suggest that the association between baby media exposure and language development depends on the approach used to analyze the data. There was no evidence that exposure to baby videos had a zero-order association with language variables, aside from a small positive relationship with expressive language in early infancy. Any effect emerged in a multivariate context that was sensitive to how variables were scored. No statistically significant negative associations were found between language variables and baby DVD/video exposure when using square-root-transformed variables. Null results also occurred when taking other steps to

Table 5
OLS Regression With All Control Variables and Race (Caucasian vs. Ethnic Minority)

Predictor variable	Receptive 6–16 mos.	Express 6–16 mos.	Language 17–27 mos.
Age	.56**/.57**	.42**/.43**	.42**/.42**
Gender (girl)	.10*/.10*	.11*/.10*	.21**/.20**
Caucasian vs. ethnic min.	.00/.00	−.09/−.07	.05/.05
Reading	.11*/.11**	.05/.05	.07/.08
Stories	.12**/.12**	.04/.03	.18**/.18**
Music	.00/.01	−.04/−.02	.09/.08
Income	.01/.01	−.10*/−.09	.06/05
Maternal education	−.06/.07	−.05/−.06	.08/.08
Baby DVDs/video	−.06/−.10*	−.01/−.05	.01/.04
Adult TV	−.04/−.04	−.08/−.06	.10/.11*
Child educational	.18*/.12*	.21**/.17**	.02/.02
Child noneducational	.04/.08	.06/.13*	.03/.02

Note. All values are standardized regression coefficients (i.e., betas). Numbers to the left of the backslash are for transformed media variables. Variables to the right are for nontransformed raw variables. Receptive = receptive language; Express = expressive language.
* $p < .05$. ** $p < .01$.

reduce the impact of outliers. Moreover, the effect sizes for even the statistically significant effects for baby video exposure were at a small or even trivial level when considering current conventions (i.e., not larger than |.10| in a standardized metric; Cohen, 1992; Ferguson, 2009). In short, our analyses cast considerable doubt on the robustness of the original conclusion regarding a large effect size estimate for exposure to baby media and language development.

It is perhaps interesting that there was a consistent negative result regarding the association between the binary no-media-exposure variable and language scores for the younger age group. Children between the ages of 6 and 16 months who were not exposed to media had lower language scores than children exposed to media. These results also held when we used the sum of all media exposure rather than a dichotomous exposure variable. Thus, the Zimmerman et al. (2007) data set could be seen as contradicting the AAP policy position on children and media exposure. In fact, the current data might be used to support a recommendation that children *should* be exposed to media, although we are cautious about any such recommendation. Moreover, exposure to educational programming appears to be positively associated with language development. Such results could be used to make a case for exposing young children to educational programming. Again, however, we think that making any recommendations to parents and policy makers using this data set would be deeply misguided.

In the end, our results fit well with recent concerns that methodological flexibility contributes to false positive results in the psychological literature (LeBel & Peters, 2011; Simmons et al., 2011). There was an exceptionally large number of control variables in the original Zimmerman et al. (2007) analysis, and our reanalysis suggests that results fluctuate depending on the covariates included in the regression models and the way different media exposure variables are treated in the analyses. In line with recent recommendations, we believe researchers must justify the inclusion of all covariates in regression analyses and explain how the presence of these variables may change the interpretation of the results (Simmons et al., 2011). Researchers should also consider whether outliers impact results. Likewise, it is important that multivariate results are not interpreted as zero-order effects. These additional steps may have tempered the original conclusions reported in Zimmerman et al. (2007).

The potential substantive results reported in this article should be replicated and more thoroughly investigated in future work. Recall that our primary objective was to illustrate concerns over researcher degrees of freedom in a contentious research area with potential implications for parents regarding baby video exposure. Our report is by no means the last word on this topic. Indeed, we strongly echo the suggestion by Zimmerman et al. (2007) regarding the value of a randomized trial in which families are encouraged to eliminate baby video exposure to provide a stronger basis for generalized causal inference about the role of baby video exposure and language development.

We should also acknowledge a few limitations of the Zimmerman et al. (2007) data set, especially in terms of the ethnic diversity of the sample. These issues impact whatever substantive conclusions readers may wish to draw from this report. Although our results held through analyses in which race/ethnicity were controlled, there were too few ethnic minorities to permit a sophisticated analysis of racial/ethnic differences concerning media exposure language development. Future research with more ethnically diverse samples is important. A further issue with the Zimmerman et al. (2007) data set concerns the specific questions asked of parents in terms of reporting media exposure on the "typical weekday" and the "typical weekend day." More sophisticated diary methods might provide a more precise accounting of media exposure and perhaps generate media exposure variables with more normal distributions. Likewise, the Zimmerman et al. (2007) approach was to ask parents how often their children watched shows within preset categories (e.g., baby videos, educational shows) rather than obtaining a comprehensive list of program titles from parents and applying a coding scheme. It is possible that a preselected group of categories may not adequately represent actual categories of media that are distinct in content and use among toddlers.

In sum, the issues surrounding exposure to media and intellectual development will continue to be a subject of debate, given the importance of early cognitive development and the widespread availability of media in contemporary society. We hope that our reanalysis of Zimmerman et al. (2007) draws attention to issues of researcher degrees of freedom and underscores the value of reanalysis for improving psychological science. The current results suggest that firm conclusions are difficult to draw from the Zimmerman et al. data set and that any effect size estimates for media exposure and language development are small (if not trivial) rather than large. We believe this message is an important corrective to the narrative that has surrounded the Zimmerman et al. study.

References

American Academy of Pediatrics. (2011). Media use by children younger than 2 years. *Pediatrics, 128,* 1040–1045. doi:10.1542/peds.2011-1753

Anderson, D. R., Huston, A. C., Schmitt, K. L., Linebarger, D. L., & Wright, J. C. (2001). Early childhood television viewing and adolescent behavior: The Recontact Study. *Monographs of the Society for Research in Child Development, 66*(1). doi:10.1111/1540-5834.00120

Azen, R., Budescu, D. V., & Reiser, B. (2001). Criticality of predictors in multiple regression. *British Journal of Mathematical and Statistical Psychology, 54,* 201–225. doi:10.1348/000711001159483

Brown, F. G. (1983). *Principles of educational and psychological testing* (3rd ed.). New York, NY: Holt, Reinhart & Winston.

Brown, J. D. (2011). The media do matter: Comment on Steinberg and Monahan (2011). *Developmental Psychology, 47,* 580–581. doi:10.1037/a0022553

Brown, J. D., L'Engle, K. L., Pardun, C. J., Guo, G., Kenneavy, K., & Jackson, C. (2006). Sexy media matter: Exposure to sexual content in music, movies, television, and magazines predicts Black and White adolescents' sexual behavior. *Pediatrics, 117,* 1018–1027. doi:10.1542/peds.2005-1406

Carta, J. J., Atwater, J. B., Greenwood, C. R., McConnell, S. R., McEvoy, M. A., & Williams, R. (2001). Effects of cumulative prenatal substance exposure and environmental risks on children's developmental trajectories. *Journal of Clinical Child Psychology, 30,* 327–337. doi:10.1207/S15374424JCCP3003_5

Christakis, D. A., Zimmerman, F., DiGiuseppe, D., & McCarty, C. (2004). Early television exposure and subsequent attentional problems in children. *Pediatrics, 113,* 708–713. doi:10.1542/peds.113.4.708

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155–159. doi:10.1037/0033-2909.112.1.155

Collins, R. L., Martino, S. C., & Elliott, M. N. (2011). Propensity scoring and the relationship between sexual media and adolescent sexual behavior: Comment on Steinberg and Monahan (2011). *Developmental Psychology, 47,* 577–579. doi:10.1037/a0022564

Courage, M., & Setliff, A. (2009). Debating the impact of television and video material on very young children: Attention, learning, and the developing brain. *Child Development Perspectives, 3*(1), 72–78. doi:10.1111/j.1750-8606.2008.00080.x

de Vries, R. E. (2011). No support for a general factor of personality in a reanalysis of Van der Linden et al. *Personality and Individual Differences, 50,* 512–516. doi:10.1016/j.paid.2010.11.020

Fanelli, D. (2010). "Positive" results increase down the hierarchy of the sciences. *PLoS ONE, 5*(4), e10068.

Fanelli, D. (2012). Negative results are disappearing from most disciplines and countries. *Scientometrics, 90,* 891–904. doi:10.1007/s11192-011-0494-7

Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics, 21,* 95–116. doi:10.1017/S0142716400001053

Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice, 40,* 532–538. doi:10.1037/a0015808

Ferguson, C. J., & Brannick, M. T. (2012). Publication bias in psychological science: Prevalence, methods for identifying and controlling, and implications for the use of meta-analyses. *Psychological Methods, 17,* 120–128. doi:10.1037/a0024445

Foster, E. M., & Watkins, S. (2010). The value of reanalysis: TV viewing and attention problems. *Child Development, 81,* 368–375. doi:10.1111/j.1467-8624.2009.01400.x

Glass, G., & Hopkins, K. (1996). *Statistical methods in education and psychology* (3rd ed.) Needham Heights, MA: Allyn & Bacon.

Greenwald, A. G., Pratkanis, A. R., Leippe, M. R., & Baumgardner, M. H. (1986). Under what conditions does theory obstruct research progress? *Psychological Review, 93,* 216–229. doi:10.1037/0033-295X.93.2.216

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine, 2*(8), e124. doi:10.1371/journal.pmed.0020124

Kashy, D. A., Donnellan, M. B., Ackerman, R. A., & Russell, D. W. (2009). Reporting and interpreting research in PSPB: Practices, principles, and pragmatics. *Personality and Social Psychology Bulletin, 35,* 1131–1142. doi:10.1177/0146167208331253

Keith, T. (2006). *Multiple regression and beyond*. Boston, MA: Pearson.

LeBel, E., & Peters, K. (2011). Fearing the future of empirical psychology: Bem's (2011) evidence of psi as a case study of deficiencies in modal research practice. *Review of General Psychology, 15,* 371–379. doi:10.1037/a0025172

Lewin, T. (2009, October 23). No Einstein in your crib? Get a refund. *The New York Times*. Retrieved from http://www.nytimes.com

Lewin, T. (2010, January 12). "Baby Einstein" founder goes to court. *The New York Times*. Retrieved from http://www.nytimes.com

Lovas, G. S. (2011). Gender and patterns of language development in mother–toddler and father–toddler dyads. *First Language, 31*(1), 83–108. doi:10.1177/0142723709359241

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7,* 19–40. doi:10.1037/1082-989X.7.1.19

Mehrens, W. A., & Lehmann, I. J. (1991). *Measurement and evaluation in education and psychology* (4th ed.). New York, NY: Holt, Reinhart & Winston.

Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science, 7,* 615–631. doi:10.1177/1745691612459058

Schmidt, M. E., Rich, M., Rifas-Shiman, S., Oken, E., & Taveras, E. (2009). Television viewing in infancy and child cognition at 3 years of age in a US cohort. *Pediatrics, 123*(3), e370–e375. doi:10.1542/peds.2008-3221

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22,* 1359–1366. doi:10.1177/0956797611417632

Simonsohn, U. (2012). *Just post it: The lesson from two cases of fabricated data detected by statistics alone*. Unpublished manuscript, University of Pennsylvania.

Steinberg, L., & Monahan, K. C. (2010). Adolescents' exposure to sexy media does not hasten the initiation of sexual intercourse. *Developmental Psychology, 47,* 562–576. doi:10.1037/a0020613

Steinberg, L., & Monahan, K. C. (2011). Premature dissemination of advice undermines our credibility as scientists: Reply to Brown (2011) and to Collins, Martino, and Elliott (2011). *Developmental Psychology, 47,* 582–584. doi:10.1037/a0022562

Stevens, J. P. (1984). Outliers and influential data points in regression analysis. *Psychological Bulletin, 95,* 334–344. doi:10.1037/0033-2909.95.2.334

Stevens, T., & Mulsow, M. (2006). There is no meaningful relationship between television exposure and symptoms of attention-deficit/hyperactivity disorder. *Pediatrics, 117,* 665–672. doi:10.1542/peds.2005-0863

University of Washington. (2007). Baby DVDs, videos, may hinder, not help, infants' language development [News release]. Retrieved December 22, 2011, from http://www.washington.edu/alumni/uwnewslinks/200709/videos.html

van der Linden, D. (2011). The general factor of personality and peer-rated social status: A rejoinder to de Vries (2011). *Personality and Individual Differences, 51,* 836–839. doi:10.1016/j.paid.2011.07.010

van der Linden, D., Scholte, R. H. J., Cillessen, A. H. N., te Nijhenhuis, J., & Segers, E. (2010). Classroom ratings of likeability and popularity are related to the Big Five and the general factor of personality. *Journal of Research in Personality, 44,* 669–672. http://dx.doi.org/10.1016/j.jrp.2010.08.007

Weigel, D. J., Martin, S. S., & Bennett, K. K. (2006). Contributions of the home literacy environment to preschool-aged children's emerging literacy and language skills. *Early Child Development and Care, 176,* 357–378. doi:10.1080/03004430500063747

Wicherts, J. M., & Bakker, M. (2012). Publish (your data) or (let the data) perish! Why not publish your data too? *Intelligence, 40,* 73–76. doi:10.1016/j.intell.2012.01.004

Wilcox, R. (2012). *Modern statistics for the social and behavioral sciences: A practical introduction*. New York, NY: CRC Press.

Zientek, L., & Thompson, B. (2006). Commonality analysis: Partitioning variance to facilitate better understanding of data. *Journal of Early Intervention, 28,* 299–307. doi:10.1177/105381510602800405

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

Received February 9, 2012
Revision received March 11, 2013
Accepted May 6, 2013 ■