

Running head: VIDEO GAME META-ANALYSIS

The Effects of Violent Video Games on Aggression:
A Meta-Analysis

John L. Sherry
Purdue University

Accepted for publication in Human Communication Research

John Sherry is an Assistant Professor in the Department of Communication at Purdue University. The author wishes to thank Bradley Greenberg for his advice and support and Charles Atkin, Bella Mody, Charles Salmon, and Glenn Sparks for feedback on earlier versions of this paper. Correspondence concerning this article should be addressed to the author at the Department of Communication, Purdue University, 1366 LAEB 2114, West Lafayette, Indiana 47907. Email may be sent to jsherry@sla.purdue.edu.

Abstract

Violent content video games such as Mortal Kombat and Doom have become very popular among children and adolescents, causing great concern for parents, teachers and policy makers. This study cumulates findings across existing empirical research on the effects of violent video games to estimate overall effect size and discern important trends and moderating variables. Results suggest there is a smaller effect of violent video games on aggression than has been found with television violence on aggression. This effect is positively associated with type of game violence and negatively related to time spent playing the games. Directions for future programmatic research on video games are outlined.

Do Violent Video Games Cause Aggression?

A Meta-Analytic Review

The arrival of each new communication technology has brought public concern about the influence of the medium on children (Wartella & Reeves, 1985). In the wake of high school tragedies at Paduchah, Kentucky and Littleton, Colorado, keen public interest has been focused on the effects of violent video games. As has been the case with violent television programs, commentators have lined up to denounce the violence found in video games while the video game industry circles the wagons in defense of its product. Unlike the television controversy, the existing social science research on the impact of video games is not nearly as compelling. Despite over thirty studies, researchers cannot agree if violent-content video games have an effect on aggression. The literature on video game effects is littered with mixed findings from studies that use a wide range of games, treatment exposure times, and subject pools, obscuring clear conclusions. Some researchers claim that playing violent video games leads to aggression (e.g., Anderson & Ford, 1986; Ballard & Wiest, 1995; Irwin & Gross, 1995; Schutte, Malouff, Post-Gordon, & Rodasta, 1988; Silvern & Williamson, 1987), while others claim there is no such relationship (e.g., Cooper & Mackie, 1986; Graybill, Kirsch, & Esselman, 1985; Graybill, Strawniak, Hunter, & O'Leary, 1987; Scott, 1995; Winkel, Novak, & Hopson, 1987). Even reviewers of the literature disagree. While Dill and Dill's (1998) interpretive literature review concluded that there is clear evidence of a causal relationship between video game play and aggression, a year later the same journal published another literature review (Griffiths, 1999) that argued methodological problems in the literature prevent a clear conclusion. Both reviewers agree that much of the research activity to date has been peripatetic, with no theory-based program of research emerging to focus systematic inquiry. The present paper attempts to bring order to the chaos and to outline future directions for research that will make best

use of available knowledge.

After a brief decline in popularity, video games have emerged again as one of the most popular media for children and adolescents. Video games are now in an estimated 80% of homes with boys aged 8-16, and video game manufacturers enjoy a larger market than the U.S. film industry (Battelle & Johnstone, 1993). A survey of middle-class adolescents showed that 88% of boys and 64% of girls reported playing video games at least one hour a week, with 29% of the boys playing 3-6 hours a week on average (Funk, 1993). The same study revealed that half of the adolescents in the sample preferred games characterized as containing human violence (17%) or fantasy violence (31.9%), with non-violent categories of sports (29.4%), general entertainment (19.7%) and educational themes (1.8%) accounting for the remainder. Fighting games (e.g., Mortal Kombat, Streetfighter, and Tekken) and first-person shooters (e.g., Quake, Doom, and Marathon) have greatly increased the level of violence over earlier games with faster and more graphic depictions of human violence. Increasingly, these games are played over the internet pitting players against each other in a virtual world. Parents and media watchdog organizations such as the National Coalition on Television Violence (NCTV, 1990) have become concerned about possible negative consequences of violent-content video games. Former United States Surgeon General C. Everett Koop went so far as to declare video games among the top three causes of family violence (Meinel, 1983). Recently, the mayor of Indianapolis spearheaded a law banning children under 18 years old from playing violent video games unless accompanied by an adult (Halladay & Wolf, 2000). Parents, educators and social activists, believing that long-term exposure to interactive violence poses a greater threat to children than other types of mediated violence, continue to protest despite enactment of a voluntary industry rating system.

One powerful way of sorting out the controversy and examining trends in existing data from

which productive, theory-building research can emerge is meta-analysis (Hunt, 1997). Essentially, meta-analysis has the potential to convert a number of small studies that may lack statistical power (due to small number of subjects) into one large study with vastly improved statistical power (the sum of subjects across all studies becomes the new N) and a wider range of variables to analyze (Hunter, Schmidt, & Jackson, 1982; Rosenthal, 1991; Rosenthal, 1995). Unlike qualitative reviews, meta-analysis is not deceived by mixed results due to problems associated with significance testing. That is, two studies resulting in the same correlation between two variables, but differing in number of subjects, may well differ on the conclusion of whether to reject the null hypothesis. Further, meta-analysis allows for statistical examination of a range of variables involved in the studies. In this case, meta-analysis allows the examination of differing game types, length of playing time, age, etc. in ways that are not feasible in the individual studies. Meta-analysis at this stage of research on video games may also serve to direct research toward important causal variables and focus inquiry in a more systematic, productive manner.

Theoretical and Methodological Issues in Video Game Research

Theoretical Concerns

Researchers have hypothesized that playing violent-content video games will result in aggression based on theories associated with violent television viewing including social learning, neo-associative networks, arousal and catharsis. Dominick (1984) points out that although video games share some characteristics with television, the medium is different in several ways that impact on theoretical mechanisms. For example, Dominick suggests that television is an essentially passive experience, whereas video games are highly active, requiring intense concentration and physical activity. Even though there is debate about the extent to which television is an active or passive medium, few would argue that television viewers are engaged to the same degree as video

game players. Television viewers can break concentration and still follow the story, but video game players cannot break concentration except during programmed rest periods. Further, video game violence tends to be highly abstract, as opposed to the realistic violence portrayed on television. Studies on television and film effects have shown that greater aggression results from the viewing of more realistic or realistically perceived violence (Atkin, 1983; Berkowitz & Alioto, 1973). Finally, the amount of exposure to violence in a given amount of time varies by individual player depending on his or her skill level. Highly skilled players may engage in more violence more frequently than players learning the game.

The most frequently cited mechanism by which the games can result in aggressive behavior is social learning theory (Alman, 1992; Brusa, 1988; Chambers & Ascione, 1987; Graybill, et al., 1985; Hoffman, 1995; Irwin & Gross, 1995; Schutte, et al., 1988; Silvern, Lang, & Williamson, 1987; Winkel, Novak, & Hopson, 1987). Social learning theory posits that behavior is learned through imitation of attractive, rewarded models (Bandura, 1994). These behaviors can become a relatively enduring part of the learner's behavioral repertoire. Proponents of SLT for video games argue that video games should have particularly powerful effects due to the high attention levels of players and the active identification of players with characters on the screen. Some video game researchers also argue that game players are rewarded directly for enacting symbolic violence, and therefore may transfer the learned aggression to the outside world.

A second explanation offered for effects of violent video games is the general arousal model (Ballard & Wiest, 1995; Brusa, 1987; Calvert & Tan, 1994; Winkel et al., 1987). Tannenbaum and Zillmann's Elementary Arousal Model (1975) holds that arousal is a heightened, non-specific drive state. As such, arousal from media communication (e.g., film or television) will heighten the already-present response an individual has. That is, "presented with a need or opportunity to

respond in some manner to a particular environmental situation, the individual will do as he would ordinarily-- but with increased energy and intensity due to the available residual arousal"

(Tannenbaum & Zillmann, 1975, p. 161, emphasis added). Supporters of the arousal model for video games suggest that highly violent video games provide the arousal that is necessary to heighten aggressive responses.

Several studies have offered a third perspective, the neo-associative networks or priming effects mechanism, as an explanation of hypothesized effects of video games on aggression (Anderson & Ford, 1986; Anderson & Morrow, 1995; Hoffman, 1995). Berkowitz's priming effects theory of media influence posits that cues from violent content may lead to aggression or hostility due to the priming of semantically related informational nodes (Berkowitz & Rogers, 1986). In the case of video games, priming effects theory would suggest that exposure to violent games will prime a series of nodes associated with violence and aggression. The priming of these violence related nodes presents the opportunity for transfer of aggressive thoughts into action (Jo & Berkowitz, 1994).

The fourth proposed mechanism posits that violent-content video games will have a positive effect on users. Seven of the located studies have addressed the issue of catharsis effects, in which violent content media are used as a safe outlet for aggressive thoughts and feelings (Calvert & Tan, 1994; Dominick, 1984; Graybill, et al., 1985; Graybill, et al., 1987; Kestenbaum & Weinstein, 1985; Silvern, et al., 1987; Silvern & Williamson, 1987). This theory, originally proposed by Feshbach (1955), holds that individuals can discharge their aggressive feelings by watching a fictional portrayal of violence in the media. While catharsis effects have long been held to be an unlikely explanation of media effects (Gunter, 1994), one survey of adolescent boys reported that they used violent-content video games to discharge aggression and manage developmental conflicts

(Kestenbaum & Weinstein, 1985). The authors argue that the interactive nature of the games allows players to act out aggression that is not allowed in the real world.

Methodological Issues

The groundwork for understanding the relationship between violent video games and aggression has been laid by a handful of researchers, but many methodological questions remain. While most of the studies use a sound experimental design, using posttest-only designs with random assignment to condition, problems remain pertaining to a set of threats to validity referred to by Cook and Campbell (1979) as "construct validity of cause and effect". Questions of construct validity of cause and effect address whether an observed causal relationship generalizes beyond the particular operationalization of the construct used in that study. Construct validity of cause relates to the independent variable and includes treatment strength and integrity, the Hawthorne effect, hypothesis guessing, and problems associated with single operationalizations of the independent variable. Construct validity of effect includes potential threats to validity related to the dependent variable including questions of validity (e.g., face, concurrent, predictive, etc.), inadequate explication, and restricted generalizability because too few facets of the dependent variable are measured.

Treatment strength of violent game stimuli has varied considerably across studies on the dimensions of type of violence portrayed and length of exposure. Early studies used games with crude graphics to symbolize violence such as Atari Boxing, in which boxers were shown from above (Graybill et al., 1985; Graybill et al. 1987); Missile Command, where dots representing bombs "blew up" lines representing missiles (Cooper & Mackie, 1986); and Berzerk, in which players shoot at "robots" shaped as squares (Graybill et al., 1987). In contrast, recent studies used such graphically violent games as Mortal Kombat (Ballard & Wiest, 1995) and Double Dragon (Irwin & Gross,

1995), in which graphically realistic, human-like characters fight to the death. Importantly for Social Learning Theory, recent games offer more realistic violence than was found in the older games. Although most realistic fighting games offer moves that cannot be replicated in the real world (e.g., jumping six feet vertically to kick someone, ripping a brain and spinal column out of another person), these games also offer easily replicated punches, kicks and elbow strikes.

In addition to differences among content and action levels of the games used in the studies, length of treatment varied from as little as 5 minutes to as much as 75 minutes of game playing. This may result in a confounding of frustration in the brief conditions because the players don't have enough time to become proficient with the game and a confounding of boredom in the longer conditions because the players may be forced to play longer than they want to. Few would argue that playing a game for 5 minutes is an equivalent experience to playing the game for 75 minutes.

Differences in outcome measures of aggression used across the studies create threats to construct validity of effect. Because studies used a variety of measures of aggression ranging from actual behavior (e.g., aggression during free play, willingness to help or hurt another) to paper-and-pencil measures of aggressive feelings (e.g., Buss-Durkee Hostility Measure, Multiple Adjective Affect Checklist), it is difficult to ascertain whether the observed findings support the popular claim that video games represent a danger to society. Do video games cause people to act aggressively or to feel aggressive or both? The choice of outcome measure has both theoretical and social consequences. Paper-and-pencil measures of hostility may reflect priming of associated nodes, but priming may not result in actual aggressive behavior. Behavioral measures such as aggression during free play and willingness to help or hurt others are more socially significant. While some researchers claim that paper-and-pencil measures are an indicator of willingness to enact aggression in the real world (Anderson & Bushman, 1997), differences uncovered by the meta-analysis could

have important implications for understanding the range of video game effects and their significance. Findings that support a strong behavioral effect would lend credence to the concerns of video game critics, whereas findings that suggest only an emotional reaction to video games would appear to be less socially threatening. Further, it is important to determine whether players may be learning an enduring behavior or simply engaging in transient reactions.

Meta-Analytic Objectives

A meta-analysis of existing studies of the effects of violent content video games on aggression will provide important information for consideration of theoretical and methodological issues leading to a systematic program of research in this area. This article will address five objectives: (a) provide a cumulative measure of effect size across all studies, (b) test for consistent differences in effect size due to levels of graphic violence in games, (c) test for the effect of length of treatment on effect size, (d) demonstrate any relationship between effect size and age of subjects, and (e) demonstrate any relationship between effect size and type of outcome measure.

Method

This section describes three phases of the meta-analysis: (a) the method of selecting and coding studies for inclusion, (b) the method of deriving effect size estimates from the studies, and (c) the procedure for statistical analysis of the effect sizes.

Study Selection and Coding. A search of literature from January 1975 to July 2000 was undertaken using on-line data bases (ERIC, PsychInfo); CD-ROM (Dissertation Abstracts International); and bound indexes (Communication Abstracts, Psychological Abstracts). Keywords varied according to the index used, but were generally either “video game” or “computer game.” Indexes were searched back to 1975, when possible, in order to pre-date research on video games. These searches resulted in excess of 900 citations, many of which were redundant across indexes.

Titles and abstracts were examined to locate studies and literature reviews that addressed the effects of video games on aggression. Any available research report fitting the narrowed description was collected and the reference sections were examined to locate additional studies. Where possible, the additional studies were added to the collection and examined for references to other studies. This process iterated until no new studies were found. During the search process, correspondence was kept with a researcher who was conducting a meta-analysis on the same topic and lists of studies were compared (M. Mittag, personal communication, August to October, 1996). Correspondence also was undertaken with other researchers in the field to locate unpublished studies.

The narrowed search resulted in 32 independent studies in which violent video game play was the independent variable and some measure of aggression was the dependent variable. Of these, one study was unavailable from the author (Favaro, 1983); three studies were excluded because they did not provide a usable control group for comparison (Alman, 1992; Anderson & Morrow, 1995; Felsten, 1995); two studies were excluded because they lacked a usable measure of aggression (Kestenbaum & Weinstein, 1985; Mehrabian & Wixen, 1986); and one study was excluded because the reported data were not interpretable (Scott, 1995). In cases where the same data sets were reported in more than one report, the available report with the most complete data was used and the others were excluded. A total of 25 studies remained for which effect size estimates could be calculated.

Study coding sheets included entries for the following characteristics: study descriptives; sample characteristics; research design; stimuli description; and results. Study descriptives include title, author(s), date of publication, and type of report (e.g., journal article, conference paper, dissertation). The earliest known publication date of a data set was used as the date of publication for the study. Sample characteristics were recorded for treatment and control conditions across the

following categories: sample size, age, % male, % Caucasian, number entering the study and number participating. Across the 25 studies, there were very few instances of subjects being dropped from the study, so this was not pursued further. Racial distribution of subjects was reported in only five of the studies. Finally, the type of sampling used to select participants was recorded.

Research design was recorded for all included studies. Six studies were cross-sectional surveys (Anderson & Dill, 1997a¹; Dominick, 1984; Fling, Smith, Rodriguez, Thorton, Atkins, & Nixon, 1992; Gibb, Faily, Lambirth, & Wilson, 1983; Lin & Lepper, 1987; Weigman & van Schie, 1994), whereas the remaining studies used an experimental design with random assignment to treatment and control conditions. When an experimental design was used, the games used in the treatment condition as well as the type of control used were recorded. The amount of playing or viewing time of treatment and control conditions were also recorded.

The coding sheet contained a blank space for the recording of results to accommodate the variety of results statistics and outcome measures across studies. Some studies used a behavioral outcome measure of aggression such as free play (Brusa, 1988; Cooper & Mackie, 1986; Irwin, 1993), help or hurt tests (Graybill, et al., 1987) or monetary deduction or donation (Chambers & Ascione, 1987; Winkel, et al., 1987). Other studies used a paper-and-pencil measure of hostile feelings such as the Multiple Affect Adjective Checklist (Anderson & Ford, 1986; Calvert & Tan, 1994; Walker, 1985), Buss-Durkee Hostility Measure (Ballard & Weist, 1995; Hoffman, 1995) or Rozenswieg Picture-Frustration (Graybill, et al., 1985; Graybill, et al., 1987).

Effect Size Estimate. Pearson's r was used as the effect size estimate for its simplicity of interpretation and flexibility across types of studies and reported results (for more on the preference of Pearson's r as an effect size estimate, see Hunter et al., 1982; Rosenthal 1991, chapter 2). In the six survey studies, reported correlations between video game use and the

hostility or aggression outcome measure were used as the effect size. For experimental studies, an effect size was calculated only when a comparison between treatment and control conditions was reported. When means, standard deviations and cell sizes were reported, Cohen's d (Cohen, 1988) was calculated and converted to Pearson's r . Cohen's d is the difference between treatment mean (\bar{y}_E) and control mean (\bar{y}_C) standardized by dividing by an estimate of the within group standard deviation (s)² and was calculated using the following formula:

$$d = \frac{\bar{y}_E - \bar{y}_C}{s} . \quad (1)$$

The within-group estimate of variance for the denominator (s^2) was calculated as outlined in Hunter, et al. (1982, p. 96-99) in order to lower the standard error of the estimate:

$$s^2 = \frac{(N_E - 1)s_E^2 + (N_C - 1)s_C^2}{N_E + N_C - 2} , \quad (2)$$

where N_E is the number of subjects in the experimental group, s_E^2 is the variance of the experimental group, N_C is number of subjects in the control group, and s_C^2 is the variance of the control group. Finally, the estimate of Cohen's d was converted to Pearson's r using Equation 3:

$$r = \frac{d}{\sqrt{d^2 + \frac{4(N - 2)}{N}}} . \quad (3)$$

In the remaining studies that did not provide means, standard deviations and cell sizes, a variety of

methods was used to extract an effect size estimate including conversion from t -tests, conversion from F with 1 degree of freedom in the numerator, and reconstruction of ANOVA tables from partial reports to calculate r^2 (for formulas, see Rosenthal, 1991 or Hunter et al., 1982).

The effect size, r , represents the relationship between video game play and some measure of aggression found in each study. In several cases, there were more than one measure of aggression in the study. Because all such measures purport to measure the same construct (aggression), this is a conceptual replication (Hendrick, 1990). Therefore, the individual study was retained as the unit of analysis and the mean of all effect sizes within each study is reported as the effect size for the study. This has the additional benefit of avoiding the problem of independence associated with meta-analysis using the hypothesis as the unit of analysis.

Often in research reports, statistics such as means, standard deviations, and t -tests are not reported if an initial test of significance fails to reject the null hypothesis. This does not mean that there is not an effect of X on Y , possibly of a fairly high magnitude if the study's N is small. In the studies examined here, the most common procedure was to not provide any additional statistical information if the study failed an initial omnibus F -test. If there are few subjects in the study, a sizable effect size could be reported as not significant. This poses a problem for the meta-analyst as to how to address non-significant results in which no additional information is provided. Rosenthal (1995) suggests that in absence of additional information, two calculations should be reported: 1) a calculation of study effect size setting the non-significant finding to zero and 2) a calculation dropping the non-significant effect size. Dropping the non-significant effect sizes ignores true findings of the study and presents an inflated effect size estimate. Setting the non-significant findings to zero retains all the findings of the study, but presents a conservative estimate if the non-significant findings were actually positive and inflates the effect size if the actual non-significant

findings were negative. Because there was a mixture of positive and negative effects in the studies included here, a mean effect size for studies reporting "non-significance" was calculated under three conditions: 1) non-significant findings set to zero; 2) non-significant findings set at the maximum possible non-significant positive effect size; and 3) non-significant findings set at the maximum possible non-significant negative effect size. Therefore a range of three possible effect sizes will be presented.

Method for Statistical Analysis of Effect Size Estimates. Hunter, Schmidt and Jackson (1982) offer a methodology for analyzing a collection of effect size estimates that is designed to account for the variance among those estimates. If all effect sizes are equivalent, there will be no variance among the effect sizes. Because this is rarely the case, it is necessary to isolate sources of unexplained variance among the studies. The unexplained variance could be due to sampling error or to moderating variables. The procedure for isolating unexplained variance can be broken down into three steps: 1) pool the effect size estimates and calculate the variance among estimates; 2) account for residual variance due to sampling error; and 3) account for residual variance due to moderating variables. Once all variance is accounted for, it is possible to draw conclusions about relationships among the variables involved in the meta-analysis.

First, the mean effect size (weighted by sample size) and variance from all studies is calculated. The weighted mean effect size is the superior estimate because it gives greater weight to large studies than to small studies (Hunter, et al., 1982) and therefore represents the best estimate of the correlation between video game play and aggression. The variance estimate represents unexplained variance associated with the mean effect size. The first possible source of unexplained variance to be investigated by meta-analysis is sampling error. The variance due to sampling error is calculated and subtracted from the unexplained variance in the average effect size. An estimate of

variance due to sampling error is calculated using the following formula:

$$s_e^2 = \frac{(1 - \bar{r}^2)^2 K}{N} \quad (4)$$

where K is the number of studies and N is the sum of subjects across studies (Hunter, et al., 1982).

If variance due to sampling error is greater than or equal to the variance in the average effect size, the variance in the average effect size can be explained as sampling error. If variance in the mean effect size remains after subtracting out variance due to sampling error, then there may be a moderator variable to explain the remaining residual variance. A moderator variable is defined as: "a variable that causes differences in the correlation between two other variables" (Hunter, et al., 1982, p. 47). In that case, a search for moderator variables is undertaken using categorical or correlational analyses.

The categorical meta-analytic procedure seeks to account for unexplained variance by subdividing estimates into theoretically sound categories; calculating the category means, variances, and sampling errors, then determining if all unexplained variance is accounted for. If there are differences in mean effect sizes between the categories and if the variance within categories is reduced to within sampling error, then differences in the categories can be said to explain the residual variance (Hunter, et al., 1982). If not, the search continues for other moderators. Correlational statistics can also be used if variables were measured at the interval level of measurement or higher and the sampling error in the effect size estimate can be controlled (Hunter, et al., 1982).

Results

A mean effect size representing the Pearson correlation between video game use and

aggression was calculated for each of the 25 studies in the meta-analysis (see Table 1). A range of effect size estimates are reported for the four studies which reported non-significant findings and provided no further information to calculate effect sizes. Table 2 contains estimates of the unweighted mean effect size ($\bar{r} = .16$); as well as three calculations of the weighted mean effect size ranging from $\bar{r} = .13$ to $\bar{r} = .16$. Next, the amount of variance in the studies that is due to sampling error was calculated and subtracted from the residual variance. In this case, the large standard deviations in the weighted effect sizes after subtracting out the effect of sampling error (range of $s = .20$ to $s = .22$) suggest that a moderating variable is present.

A significance test for heterogeneity of effect size among studies is also available and is distributed as χ^2 with $K-1$ degrees of freedom where K is the number of studies (Rosenthal, 1991, p. 74). The formula is:

$$\chi^2 = (N_j - 3) \left(z_{rj} - \bar{z}_r \right)^2 \quad (5)$$

where z_{rj} is the Fisher z_r corresponding to any \bar{r} and \bar{z}_r is the weighted mean z_r . Performing this test on the 25 effect sizes results in $\chi^2(24) = 151.18, p < .001$. Failing to reject the null in this case confirms a high probability of moderating variables.

Categorical Analyses-- Methodology. A series of categorical analyses were undertaken to determine whether the residual unexplained variance was due to differences in methodology including: survey vs. experiment and type of outcome aggression measure. In the case of study design, there is a small difference in effect size means between experimental ($\bar{r} = .11$) and survey studies ($\bar{r} = .16$). Additionally, there is some remaining unexplained variance (experiment $s = .01$ and survey $s = .03$).

Next, a categorical analysis was conducted comparing type of outcome measure used. Studies were categorized according to whether the outcome measure was behavioral in nature (e.g., free play, reward or punishment) or paper-and-pencil type (e.g., Buss-Durkee, Multiple Adjective Affect Checklist). Again, the means differ (behavioral measure, $\bar{r} = .09$; paper-and-pencil measure, $\bar{r} = .19$) and there is residual unexplained variance remaining in the paper and pencil measure (behavioral measure, $\tau^2 = .00$; paper and pencil measure, $\tau^2 = .04$).

Correlational Analysis-- Theoretical Moderators. A wide range of theoretically significant variables may moderate the effect of violent video games on aggression including: age of the subjects; length of exposure to the game; and type of game violence. Age of subjects and length of exposure are ratio level data and can be examined correlationally. An imprecise interval level proxy for game violence is year of study, because games used in recent studies are more violent and realistic than those used in early studies (e.g., Mortal Kombat vs. Space Invaders). Therefore, the earliest year that data were reported is also examined correlationally.

Effect size is moderately correlated with the year in which the study was performed ($r = .39$) suggesting that effect sizes have increased over time (see Table 3). Playing time was not correlated with effect size ($r = .008$). A correlation of $r = .20$ was obtained between effect size and subject age suggesting that older subjects were affected by the video games more than younger subjects. Because the moderator variables were correlated with one another (range $r = .35$ to $r = .36$), multiple regression³ was run to control for the effect of moderators on each other (e.g., suppression). The regression equation resulted in a multiple correlation coefficient of $R = .45$ ($R^2 = .21$). After controlling for the effects of the other moderators, the year in which the study was performed ($\beta = .33$) remained the strongest positive predictor of effect size, while playing time emerged as a negative predictor of effect size ($\beta = -.19$). Age of the subjects continued to be a positive predictor

of effect size ($d = .25$).

Because the year of the study (the proxy for game violence level) was a positive predictor of effect size, a categorical analysis of game traits was undertaken. For this analysis, the six surveys were dropped because they do not test a specific game. The Graybill et al. (1987) study was dropped because the reporting did not allow analysis by game, the Walker (1985) study dropped out because there was a confounding anger manipulation that cannot be statistically removed, and the Winkel, et al. (1987) study dropped out because the game used was created by the researcher and therefore there is no way to categorize it. This left sixteen studies in the meta-analysis.

A useful, tested typology for categorizing video games was provided by Funk (1993). Her system categorizes video game themes as: human violence; fantasy violence; sports; general entertainment; and educational. Using a panel of 38 children, ages 8 to 17, to rate 652 games, she attained an intercoder reliability of 94% suggesting that the categories were well conceived and consistent with the impressions of the children. For the purpose of this study, the categories of human violence, fantasy violence and sports were used. Human violence was operationalized by Funk as, "A human character must fight or destroy things and avoid being killed or destroyed while trying to reach a goal, rescue someone, or escape from something" (p. 87). Fantasy violence was defined as, "A cartoon or other fantasy character must fight or destroy things and avoid being killed or destroyed while trying to reach a goal, rescue someone, or escape from something" (p. 87). The sports category is made up of games in which, "The main action is any kind of sports, including racing and wrestling" (p.87). Two outside coders were used to assign the stimulus games to Funk's categories based on the definitions Funk provides, the game descriptions in the studies, and the coders' knowledge of the games. There were no disagreements between the coders. Weighted means in the human violence category ($\bar{r} = .15$) is equal to that of fantasy violence ($\bar{r} = .15$), with

the sports category smaller than both ($\bar{r} = .08$). The standard deviation for fantasy violence and sports is zero after subtracting out sampling error, with a small amount of residual variance in the human violence category ($\sigma^2 = .03$).

Discussion

This meta-analysis set out to further clarify methodological and theoretical issues in the literature on the effects of violent video games on aggression which will lead to more focused programmatic research. Specifically, the meta-analysis was designed to: (a) provide a cumulative measure of effect size across all studies, (b) test for consistent differences in effect size due to levels of graphic violence in games, (c) test for the effect of length of treatment on effect size, (d) demonstrate any relationship between effect size and age of subjects, and (e) demonstrate any relationship between effect size and type of outcome measure.

Overall Effect

By combining the results of a number of studies and therefore increasing the number of subjects, meta-analysis provides a robust estimate of the overall relationship between two variables by reducing the probability of alpha error and increasing statistical power ($1 - \alpha$). The overall estimate of the correlation between video game play and aggression in this meta-analysis ($r = .15$, $n = 2722$) is associated with a Type I error probability of $\alpha < .0001$ and a Type II error probability of $\beta < .05$. Cohen (1988) provides a guideline for understanding magnitude of effect size in terms of small ($d = .20$), medium ($d = .50$) and large ($d = .80$) effects. Converting the overall effect size of video game play on aggression into Cohen's d metric reveals a small effect size of $d = .30$. For comparison, Paik and Comstock (1994) arrived at an effect size estimate for the effect of television violence on aggression nearing the large effect range ($d = .65$). Overall, this analysis suggests that there is a correlation between video game play and aggression, but that relationship is smaller than

that found for television.

Graphic Violence Level

More recent games which contain both fantasy and human characters engaging in violence registered greater effect sizes than games in which the portrayed violence was sports related. This pattern of effect sizes suggests that players may be reacting to the cultural meaning of the violence that differs by game type, as has also been found in studies of violent television (Paik & Comstock, 1994). What is it about the fantasy- and human-violence games that resulted in greater aggression? In order to better understand this relationship, a more focused analysis of game elements is necessary. Perhaps the fantasy- and human-violence games contained a greater amount of action, thus raising non-specific arousal. They may also be more graphic, thus providing a more powerful prime for associated networks than more abstract images found in the older sports games. Here, a comparison between newer, more graphic boxing or sports games and the other types of violent games would be useful.

There are two important caveats to consider in interpreting these findings. First, there were only three sports games in the meta-analysis, as opposed to seven in each of the other categories. A greater number of studies would provide more confidence in the conclusions. Secondly, it must be acknowledged that game type is confounded with the year of the study. Human-violence games were not tested until the most recent studies due to availability of the games. This is important because human violence games were tested on a generational cohort that had a different experience of video games from earlier cohorts due to the proliferation of faster and more violently graphic games in the 1990's. Ultimately, the meta-analysis cannot account for these generational differences.

Length of Treatment (Playing Time)

One of the most intriguing findings of the meta-analysis is that effect size was negatively related to playing time, when controlling for age of players and year of study. Examining the partial plot of playing time on effect size controlling for age of subjects and year of study, much of this relationship is anchored by two studies, with the remainder of the studies grouped in the center and supporting the general trend. The two studies anchoring the negative slope are interesting because they were performed using the same game (Mortal Kombat), the same type of subject pool (university undergraduates), and the same outcome measure; but with widely differing playing times of 10 minutes and 75 minutes (Ballard & Wiest, 1995 and Hoffman, 1995, respectively). Ballard and Wiest found an effect size of $r = .90$ using the Buss-Durkee measure of aggression, while Hoffman found an effect size of only $r = .05$ using the same scale. The results suggest that playing even the most violent of games for extended times may not increase aggression.

Methodologically, the results suggest that the studies may be measuring an initial arousal effect that may fall off dramatically after extended play. While the initial experience of the game might be highly stimulating, after 75 minutes arousal may be replaced by fatigue or boredom from repeating programmed sequences. If arousal does drop after long periods of game play, arousal theory would predict that the aggressive reaction would also be reduced. Future research needs to carefully control or assess arousal level to better understand the role of arousal in video games effects. This can be done in a number of ways. Experiments will need to include direct measures of arousal with a number of different types of games and playing times to trace out the trajectory of the effect. Arousal may also be a function of user context. For example, a player may be more aroused when playing against another human being rather than the machine. Finally, it is also possible that players may be using the games to equilibrate their arousal levels. This suggests a drive-reduction or

catharsis hypothesis. Because none of the studies here directly test catharsis, a strong direct test of the catharsis hypothesis needs to be devised. To properly test the hypothesis, individuals must be in an angered state prior to the video game manipulation. Only then can we test to see if anger actually decreases as a result of video game use. Further, individuals should be given the option of diffusing their anger in a number of different ways (e.g., various types of games, books, movies, etc.). Finally, the individuals will need to end the session by choice. That is, they must be allowed to indicate whether the anger diffusion mechanism they used worked for them.

The results also have socially significant implications. Children and adolescents playing games in long stretches may transfer less aggression from the game playing situation to the external world than those playing for brief periods. Parents' intuitive reaction to limit playing time may actually be counter-productive, pulling the child from the game at a time when the largest aggressive effects are likely. The shrinking effect size may reflect boredom with the game, catharsis, or desensitization to the violence in the game. Studies that manipulate playing time while controlling desensitization, arousal and attention need to be conducted in order to sort out answers to these questions.

Methodological Issues-- Outcome Measure and Research Design

The average effect size for paper-and-pencil measures, which measure more attitudinal or affective aggression, was somewhat larger than that for behavioral measures. This result is to be expected because we may feel hostile, but may not act out our aggression due to cultural constraints on aggressive behavior. While this finding is expected, it is still important because it provides evidence that playing violent video games is not a powerful enough prime on average to overcome social sanctions against aggressive behaviors. Future researchers may want to focus attention on individuals who are most willing to override social sanctions after playing games and enact

behavioral aggression. Who are these individuals and why are they more willing to aggress after the video game stimulus?

Consistent with the outcome measure finding, surveys had slightly larger effect sizes than experimental studies. Surveys used paper-and-pencil measures of such constructs as attitudes about aggression or autobiographical memory of past aggressive behavior. While several of the experiments used paper-and-pencil measures of aggression, others used behavioral measures. It may be that differences in outcome measures are reflected in differences by methodology.

Conspicuously absent from the video game research are other designs used in the study of television violence such as longitudinal designs and field experiments. These types of research designs are more complex and expensive to undertake, so their absence may merely reflect the fact that video game research has only recently begun. However, these designs often provide the greatest ecological validity and allow researchers to make stronger predictions of social significance. These types of designs are crucial to understanding some of the relationships whose explanation is missing in the existing studies. For example, there is a small indication that the effect size increases as the subjects get older, controlling for playing time and game type (year of study). This finding seems counterintuitive-- we would expect younger children to be more vulnerable to the effects of video games. A longitudinal design would be helpful in assessing the effects of the games throughout the life span. A cohort study would also be helpful in settling the priming effects vs. social learning debate, as social learning predicts relatively enduring effects while priming effects predicts short-term effects. How long do the effects last? Are they translated into the real world? Field experiments would be helpful in assessing the social significance of the video game effects.

Conclusions and Directions for New Research

While meta-analysis is a useful tool for estimating true population effect sizes and isolating

trends in the literature, it is unable to establish causal relationships between variables under study. In addition, the small number of studies and the use of sub-analyses presented here enhance the possibility of capitalization on chance. Nevertheless, this analysis sheds light on important relationships and suggests paths for future, more programmatic research. Overall, the meta-analysis suggests several conclusions. First, there is a small effect of video game play on aggression, and the effect is smaller than the effect of violent television on aggression. Next, within the range of games studied, the type of violence contained in the games is a predictor of aggression, with human and fantasy violence being associated with stronger effects than sports violence. Third, there is a trend suggesting that longer playing times result in less aggression. Finally, the studies provide mixed support for various theories of aggression due to television viewing.

Further research is needed to explore the relationships among a variety of variables implicated in the potential violent video game and aggression connection. Relationships requiring further exploration address treatment strength of the violent video game induction, including a closer analysis of the effects of different game attributes (e.g. action, graphicness, difficulty, human qualities) and a more complete range of game exposure times. An early video game experiment that provides a prototype for this type of research (Anderson & Ford, 1986) was a two-stage design in which games were first rated by a pool of subjects on violent content and graphics, action, lack of pauses, difficulty, enjoyment, and frustration. A separate pool of subjects then took part in an experiment using games that differed by level of violence (mildly aggressive vs. highly aggressive), but were relatively homogenous on the other dimensions. A similar experiment using a broad range of games varying on theoretically salient dimensions will reveal game characteristics which best predict aggressive outcomes. Information gained from such a study can be applied to an additional experiment in which the most important dimensions are examined with varying playing times.

In order to understand the social significance of video games, we must understand the mechanisms by which the aggressive behavior is caused. If observed aggression after playing video games is due to arousal and priming of associative networks, the effects of violent video games should be limited in temporal duration and are of greatest concern primarily where the game playing takes place (home or arcades). Parents can expect their children to be aroused in the short term after game playing and can make rules accordingly. Arcade owners are confronted with patrons who are primed to act aggressively, thus exacerbating some of the social problems traditionally associated with arcades. New research needs to probe the mechanisms of these theories, not simply demonstrate behavioral effects.

In addition, researchers in this area will need to develop new theories that acknowledge experiential and social differences between video game use and television viewing. These theories will need to account for differences in engagement, concentration, modeled vs. actual rewards, and time spent with games. Further, such theories will need to acknowledge the social settings in which the games are played and the meaning that accompanies the games. Video games are often played in groups with social status accorded to those who score highest or beat their friends. While stories (e.g., folk narrative to television) have been used to educate people about socially appropriate behavior since the beginning of oral culture, games (e.g., chess to video games) have served a different purpose. What is the social purpose of game play and how does it relate to the video game experience? These questions need to be addressed at a theoretical level before we can adequately explain the effects of violent video games.

References

- Alman, R. E. (1992). Video games: Interaction vs. observation as sources of social learning. Unpublished master's thesis, Michigan State University, East Lansing.
- Anderson, C.A., & Bushman, B.J. (1997). External validity of "trivial" experiments: The case of laboratory aggression. Review of General Psychology, 1, 19-41.
- Anderson, C. A., & Dill, K. E. (2000). Video games and aggressive thoughts, feelings, and behavior in the laboratory and in life. Journal of Personality and Social Psychology, 78, 772-790.
- Anderson, C. A., & Ford, C. M. (1986). Affect of the game player: Short-term effects of highly and mildly aggressive video games. Personality and Social Psychology Bulletin, 12, 390-402.
- Atkin, C. (1983). Effects of realistic TV violence vs. fictional violence on aggression. Journalism Quarterly, 60, 615-621.
- Ballard, M. E., & Lineberger, R. (1999). Video game violence and confederate gender: Effects on reward and punishment given by college males. Sex Roles, 41, 541-558.
- Ballard, M. E., & Wiest, J. R. (1995, March). Mortal Kombat: The effects of violent video technology on males' hostility and cardiovascular responding. Paper presented at the Biennial Meeting of the Society for Research in Child Development, Indianapolis, IN.
- Bandura, A. (1994). The social cognitive theory of mass communication. In J. Bryant & D. Zillmann (Eds.), Media effects: Advances in theory and research (pp. 61-90). Hillsdale, NJ: Erlbaum.
- Battelle, J., & Johnstone, B. (1993, December). Seizing the next level: Sega's plan for world domination. Wired, pp. 73-78, 128-131.
- Berkowitz, L., & Alioto, J. (1973). The meaning of an observed event as a determinant of its aggressive consequences. Journal of Personality and Social Psychology, 28, 206-217.

- Berkowitz, L., & Rogers, K. H. (1986). A priming effect analysis of media influences. In J. Bryant & D. Zillmann (Eds.), Perspectives on media effects (pp. 57-81). Hillsdale, NJ: Erlbaum.
- Brusa, J.A. (1988). Effects of video game playing on children's social behavior (aggression, cooperation) (Doctoral dissertation, DePaul University, 1987). Dissertation Abstracts International-B, 48/10, 3127.
- Calvert, S., & Tan, S. L. (1994). Impact of virtual reality on young adults' physiological arousal and aggressive thoughts: Interaction versus observation. Journal of Applied Developmental Psychology, 15, 125-139.
- Chambers, J. H., & Ascione, F. R. (1987). The effects of prosocial and aggressive video games on children's donating and helping. Journal of Genetic Psychology, 148, 499-505.
- Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cook, T. D., & Campbell, D. T. (1979). Quasi-experimentation: Design and analysis issues for field research. Chicago: Rand McNally.
- Cooper, J., & Mackie, D. (1986). Video games and aggression in children. Journal of Applied Social Psychology, 16, 726-744.
- Dill, K. E. (1997). Violent video game and trait aggression effects on aggressive behavior, thoughts, and feelings, delinquency, and world view. Dissertation Abstracts International-B, 59/07, (UMI No. 9841277)
- Dill, K. E., & Dill, J. C. (1998). Video game violence: A review of the empirical literature. Aggression & Violent Behavior, 3, 407-428.
- Dominick, J. R. (1984). Videogames, television violence, and aggression in teenagers. Journal of Communication, 34, 136-147.

Feshbach, S. (1955). The drive-reducing function of fantasy behavior. Journal of Abnormal and Social Psychology, 50, 3-11.

Funk, J. (1993). Reevaluating the impact of video games. Clinical Pediatrics, 32, 86-90.

Graybill, D., Kirsch, J., & Esselman, E. (1985). Effects of playing violent versus nonviolent video games on the aggressive ideation of aggressive and nonaggressive children. Child Study Journal, 15, 199-205.

Graybill, D., Strawniak, M., Hunter, T., & O'Leary, M. (1987). Effects of playing versus observing violent versus nonviolent video games on children's aggression. Psychology: A Quarterly Journal of Human Behavior, 24(3), 1-8.

Griffiths, M. (1999). Violent video games and aggression: A review of the literature. Aggression & Violent Behavior, 4, 203-212.

Gunter, B. (1994) The question of media violence. In J. Bryant & D. Zillmann (Eds.), Media effects: Advances in theory and research (pp. 163-212). Hillsdale, NJ: Erlbaum.

Halladay, J., & Wolf, R. (2000, July 18). Indianapolis OKs restrictions on violent video game usage. USA Today, A5.

Hendrick, C. (1990). Replications, strict replications, and conceptual replications: Are they important? In Neuliep, J. W. (Ed.) Handbook of replication research in the behavioral and social sciences. [Special Issue.] Journal of Social Behavior and Personality, 5(4), 41-49.

Hoffman, K. (1995). Effects of playing versus witnessing video game violence on attitudes toward aggression and acceptance of violence as a means of conflict resolution (Doctoral dissertation, University of Alabama, 1994). Dissertation Abstracts International, 56/03, 747.

Hunt, M. M. (1997). How science takes stock : The story of meta-analysis. New York : Russell Sage Foundation.

Hunter, J. E., Schmidt, F. L., & Jackson, G. B. (1982). Meta-analysis: Cumulating research findings across studies. Beverly Hills, CA: Sage.

Irwin, A. R., & Gross, A.M. (1995). Cognitive tempo, violent video games, and aggressive behavior in young boys. Journal of Family Violence, 10, 337-350.

Jo, E., & Berkowitz, L. (1994). A priming effect analysis of media influences: An update. In J. Bryant & D. Zillmann (Eds.), Media effects: Advances in theory and research (pp. 43-60). Hillsdale, NJ: Erlbaum.

Kestenbaum, G. I., & Weinstein, L. (1985). Personality, psychopathology and developmental issues in male adolescent video game use. Journal of the American Academy of Child Psychiatry, 24, 329-337.

Kirsh, S. J. (1998). Seeing the world through Mortal Kombat-colored glasses: Violent video games and the development of a short-term hostile attribution bias. Childhood, 5, 177-184.

Meinel, C. (1983, May/June). Will Pac-Man consume our nation's youth? Technology Review, pp. 10, 11, 28.

National Coalition of Television Violence. (1990). Nintendo tainted by extreme violence. NCTV News, 11, 3-4.

Paik, H., & Comstock, G. (1994). The effects of television violence on antisocial behavior: A meta-analysis. Communication Research, 21, 516-546.

Rosenthal, R. (1995). Writing meta-analytic reviews. Psychological Bulletin, 118, 183-192.

Rosenthal, R. (1991). Meta-analytic procedures for social research. Newbury Park, CA: Sage.

Schutte, N., Malouff, J., Post-Gordon, J., & Rodasta, A. (1988). Effects of playing video games on children's aggressive and other behaviors. Journal of Applied Social Psychology, 18, 451-

456.

Scott, D. (1995). The effect of video games on feelings of aggression. Journal of Psychology, 129, 121-132.

Silvern, S. B., Lang, M. K., & Williamson, P. A. (1987). Social impact of video game play. Meaningful Play, Playful Meaning: Proceedings of the 11th Annual Meeting of the Association for the Anthropological Study of Play. Champaign, IL: Human Kinetics.

Silvern, S., & Williamson, P. (1987). The effects of video game play on young children's aggression, fantasy, and prosocial behavior. Journal of Developmental Psychology, 8, 449-458.

Tannenbaum, P. H., & Zillmann, D. (1975). Emotional arousal in the facilitation of aggression through communication. In L. Berkowitz (Ed.), Advances in experimental social psychology: Vol. 8 (pp. 149-192). New York: Academic Press.

Walker, M. R. (1985). The effects of video games and TV/film violence on subsequent aggression in male adolescents (Doctoral Dissertation, University of Southern Mississippi, 1985). Dissertation Abstracts International, 46, 2082.

Wartella, E., & Reeves, B. (1985). Historical trends in research on children and the media: 1900-1960. Journal of Communication, 35, 118-133.

Wiegman, O., & van Schie, E. G. M. (1998). Video game playing and its relations with aggressive and prosocial behaviour. British Journal of Social Psychology, 37, 367-378.

Winkel, M., Novak, D., & Hopson, H. (1987). Personality factors, subject gender and the effects of aggressive video games on aggression in adolescents. Journal of Research in Personality, 21, 211-223.

Footnotes

¹The Anderson and Dill (2000) article is based on data from Dill's (1997) dissertation and contains two separate studies for which effect sizes were calculated. In this paper, the survey study is designated "Anderson & Dill(a)" and the experimental study is designated "Anderson & Dill(b)" even though they are both from the same article.

² Variances of conceptually replicated effect sizes were found to be equal using Levene's test for homogeneity of variance.

³Regression residuals were normal and there was no evidence of collinearity in the regression coefficients based on examination of tolerances. A search for outliers was conducted using Mahalanobis's Distance and none were found.

Table 1

Effect Size Estimates for the 20 Included Studies

Study	Age	Mean effect size	Maximum negative	Maximum positive	N
Anderson & Dill (a)	18.5	.46			227
Anderson & Dill (b)	18.5	.07			196
Anderson & Ford	18-22	.36			60
Austin	19.7	.06			102
Ballard & Lineberger	21	.04			119
Ballard & Wiest	18-20	.89			30
Brusa	6	.08			32
Calvert & Tan	20.5	.34			36
Chambers and Ascione	8 to 15	.13			160
Cooper and Mackie ^a	9 to 11	.05	-.04	.13	84
Dominick	15-16	.07			250
Fling	11 to 17	.23			153
Gibb	12 to 34	-.05			140
Graybill, et al. 1985	2nd, 4th, 6th	.01			116
Graybill, et al. 1987	2nd, 4th, 6th	.00			146
Hoffman ^a	18-22	.06	-.04	.16	64
Irwin	7 to 8	.32			60
Kirsh	9.8	.17			52
Lin and Lepper	4th, 5th, 6th	.19			210

Schutte et al.	5 to 7	.17			31
Silvern & Williamson	4 to 6	.16			14
Silvern, Lang & Williamson ^a	6 to 9	.05	.00	.10	46
Walker ^a	18-19	.00	-.28	.28	56
Weigman & van Schie	11.5	.05			278
Winkel, et al.	8th grade	.06			60

Note. ^a Mean effect size with “non-significant” findings set to 0 is in the first column.

Table 2

Mean effect sizes and residual variance

<u>Criterion</u>	<u>Mean r</u>	<u>sd</u>
Unweighted effect sizes	.16	.20
Weighted effect sizes	.15	.21 ^a
Weighted effect size (max. positive)	.16	.20 ^a
Weighted effect size (max. negative)	.13	.22 ^a

Note. ^a After correction for sampling error

N= 1716

Table 3

Correlation between moderator variables and effect size

	Subject age	Playing time	Year	Effect size
Subject age	1.00			
Playing time	.35	1.00		
Year	.36	.35	1.00	
Effect size	.20	.01	.39	1.00