

Advertisement

CALL FOR PAPERS



SCIENCE PARTNER JOURNALS




Research An Open Access Journal
OFFICIAL JOURNAL OF CAST

Log in | My account | Contact us

Become a member Renew my subscription | Sign up for newsletters



Meta-analyses have failed to settle the question of whether violent video games such as *Grand Theft Auto* cause aggression. CATE GILLON/GETTY IMAGES

f 253



1



Meta-analyses were supposed to end scientific debates. Often, they only cause more controversy

By **Jop de Vrieze** | Sep. 18, 2018 , 4:15 PM

After Nikolas Cruz killed 17 students and teachers and wounded 17 others early this year at Marjory Stoneman Douglas High School in Parkland, Florida, President Donald Trump had a theory about the underlying causes. "I'm hearing more and more people say the level of violence on video games is really shaping young people's thoughts," he tweeted.

Related stories on metaresearch

[‘Journalologists’ use scientific methods to study academic publishing. Is their work improving science?](#)



He wasn't the only one to make the connection. Claims about a link to violence in movies and games surface after almost every mass shooting, and it's easy to see why, once you watch someone kill hundreds of anonymous enemies with automatic weapons and trench bombs in the bestselling game *Call of Duty*, or murder innocent drivers in the wildly popular *Grand Theft Auto*. Cruz reportedly loved such games: "It was kill, kill, kill, blow up something, and kill some more, all day," a former neighbor told *The Miami Herald*.

Yet the hundreds of scientific studies that have explored whether media violence can lead to aggressive thoughts and actions have produced conflicting results. That's why scientists have resorted to meta-analyses, studies that collect all the evidence about a scientific question, weigh it impartially, and declare a winner.

SIGN UP FOR OUR DAILY NEWSLETTER

Get more great content like this delivered right to you!

Country *	▼
Email Address *	Sign Up

[Click to view the privacy policy.](#)

Required fields are indicated by an asterisk (*)

In 2009, Christopher Ferguson and John Kilburn of Texas A&M International University in Laredo published **a major meta-analysis** in the *Journal of Pediatrics* that dismissed the link between media violence and aggression. A year later, however, a team led by psychologists Brad Bushman, then at the University of Michigan in Ann Arbor, and Craig Anderson of Iowa State University in Ames published **an even bigger analysis** in the *Psychological Bulletin* that found the opposite: The evidence "strongly suggests" exposure to violent video games is a causal factor for increased aggressive feelings and behavior, they wrote.

Researchers on both sides of the issue are wedded to their views. Ferguson, a gamer himself, has written, "The exaggerated focus on violent video games distracts society from much more important causes of aggression." Bushman, for his part, has consistently found a link between video games and violence in his own studies, which span 2 decades. (He retracted two papers recently, after critics exposed irregularities in the data, and a third because of alleged self-plagiarism.) Since their meta-analyses were published, Bushman, now at The Ohio State University in Columbus, and Ferguson have fought an increasingly fierce and sometimes personal battle, but the question

remains unresolved. Several other meta-analyses have been published, but none has settled the matter.

Similar fights are playing out in other fields. Although the number of meta-analyses has exploded, many don't bring clarity—whether it's on the effect of “positive parenting,” the relation between antidepressants and suicide, or the health benefits of organic produce.

“ Meta-analyses were thought to be debate enders, but now we know they rarely are. ”

Christopher Ferguson, Texas A&M International University

One reason is that, although the basic rules of the meta-analysis are simple, researchers must make many choices along the way, allowing conscious or unconscious biases to creep in. In the case of media violence, for instance, the groups dealt in different ways with the problem that many studies aren't published, and they applied different quality criteria in choosing the studies to be included.

“Meta-analyses were thought to be debate enders, but now we know they rarely are,” Ferguson says. “They should be regarded as an argument, not a fact.” It's a paradox, says Jacob Stegenga, a philosopher of science at University of Cambridge in the United Kingdom: “When the evidence points clearly in one direction, there is little need for a meta-analysis. When it doesn't, a meta-analysis is unlikely to give the final answer.”

Still, some metaresearchers say there are ways to do better. The movement in some fields to ensure that all studies are published, whatever their outcome, will help ensure that meta-analyses can take the full range of evidence into account. Meta-analyses themselves can be done better. And some say researchers at odds over issues such as violence in the media should sign a truce, join hands, and design a meta-analysis that everybody can agree on.

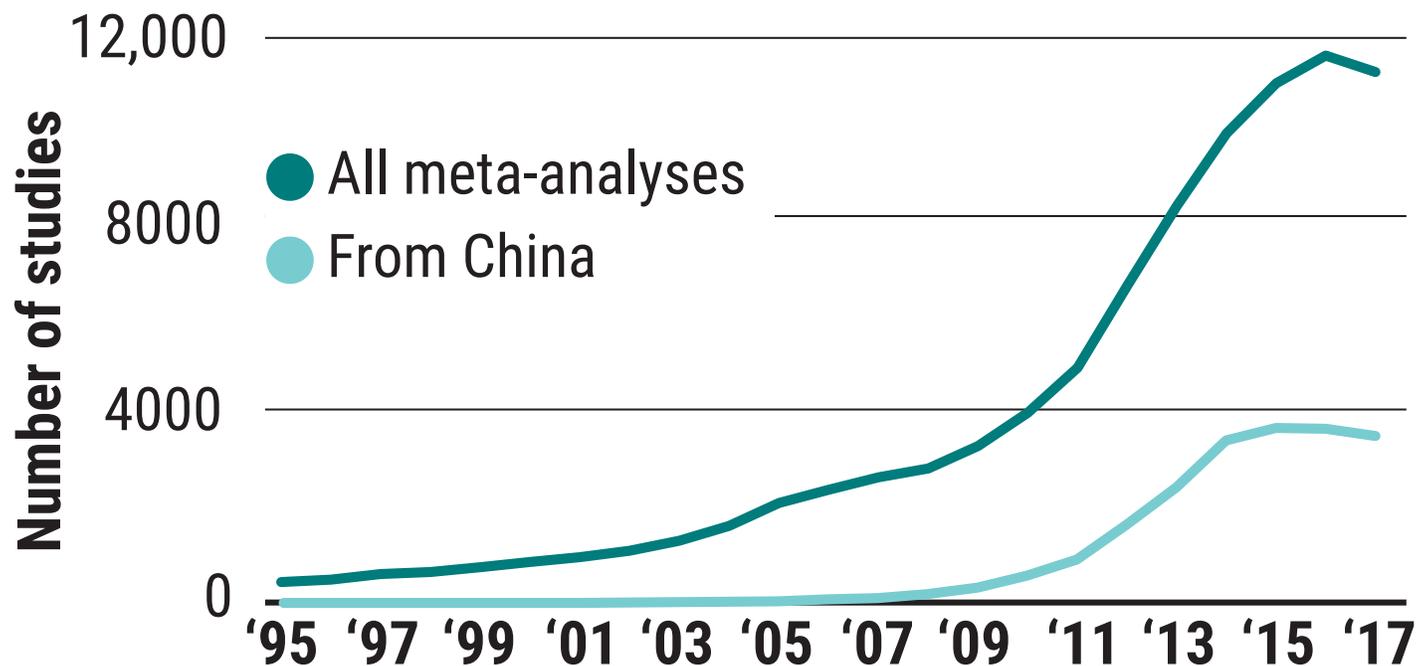
The term meta-analysis was **coined** in 1976 by statistician Gene Glass of the University of Colorado in Boulder, who described it as “an analysis of analyses.” Glass, who worked in education psychology, had undergone a psychoanalytic treatment and found it to work very well; he was annoyed by critics of psychoanalysis, including Hans Eysenck, a famous psychologist at King's College London, who Glass said was cherry picking studies to show that psychoanalysis wasn't effective, whereas behavior therapy was.

At the time, most literature reviews took a narrative approach; a prominent scientist would walk the reader through their selection of available studies and draw conclusions at the end. Glass introduced the concept of a systematic review, in which the literature is scoured using predefined search and selection criteria. Papers that don't meet those criteria are tossed out; the remaining ones are screened and the key data are extracted. If the process yields enough reasonably similar

quantitative data, the reviewer can do the actual meta-analysis, a combined analysis in which the studies' effect sizes are weighed.

Eager for answers

Virtually unknown until the 1990s, meta-analyses have recently become increasingly popular. More than 11,000 were published last year, one-third of them by authors from China.



J. YOU/SCIENCE

When Glass did this for studies of psycho-analysis, the result bore out his personal experience of its efficacy. "All the behavior therapists were outraged and all the Freudians said they knew it all along," says Glass, now 77 and retired. Eysenck was not impressed; he called meta-analyses "**an exercise in mega-silliness**" and "abuse of research integration."

Yet over the decades, the tandem of systematic review and meta-analysis has become widely accepted as a standardized, less biased way to weigh the evidence; it now guides thousands of treatment guidelines and social policies. Much of the respect it has earned reflects the work of **Cochrane**, a multinational organization headquartered in London that conducts systematic reviews of health care interventions and diagnostic tests, which are published in the Cochrane Library. (**Cochrane was plunged into crisis last week after its Governing Board voted to expel Peter Gøtzsche, a prominent member.**) The Oslo-based **Campbell Collaboration** produces similar reviews for the social sciences. Both groups follow strict **protocols** and attempt to team up experts in the

issue under study—say, cardiologists when the meta-analysis is about a heart drug—with methodological experts.

Today, meta-analyses are a growth industry. Their number has shot up from fewer than 1000 in the year 2000 to some 11,000 last year. The increase was most pronounced in China, which now accounts for about one-third of all meta-analyses. Metaresearcher John Ioannidis of Stanford University in Palo Alto, California, **has suggested** meta-analyses may be so popular because they can be done with little or no money, are publishable in high-impact journals, and are often cited.

Yet they are less authoritative than they seem, in part because of what methodologists call “many researcher degrees of freedom.” “Scientists have to make several decisions and judgment calls that influence the outcome of a meta-analysis,” says Jos Kleijnen, founder of the company **Kleijnen Systematic Reviews** in Escrick, U.K. They can include or exclude certain study types, limit the time period, include only English-language publications or peer-reviewed papers, and apply strict or loose study quality criteria, for instance. “All these steps have a certain degree of subjectivity,” Kleijnen says. “Anyone who wants to manipulate has endless possibilities.”

His company analyzed 7212 systematic reviews and concluded that when Cochrane reviews were set aside, only 27% of the meta-analyses had a “low risk of bias.” Among Cochrane reviews, 87% were at low risk of bias.

A good meta-analysis starts with clear criteria for study inclusion and exclusion, says statistician Robbie van Aert, a postdoctoral researcher at Tilburg University in the Netherlands. “If you do it after you have collected the studies, you can get almost any result you want.” But bias can occur even when inclusion criteria are chosen beforehand; because experts in a field already know the relevant literature, they can consciously or unconsciously adjust the criteria to include studies they like or exclude ones they distrust, Van Aert says.

Money is one potential source of bias. It may not affect the actual results the authors produce, but it appears to affect their spin when they draw their conclusions. In 2006, for instance, the Nordic Cochrane Centre in Copenhagen **compared** Cochrane meta-analyses of drug efficacy, which are never funded by the industry, with those produced by other groups. It found that seven industry-funded reviews all had conclusions that recommended the drug without reservations; none of the Cochrane analyses of the same drugs did. Industry-funded systematic reviews also tended to be less transparent. Ioannidis **found in a 2016 review** that industry-sponsored meta-analyses of antidepressant efficacy almost never mentioned caveats about the drugs in their abstracts. “This is a clear example of an area where meta-analyses are emerging as a powerful marketing tool,” he wrote.



Meta-analyses disagree about the effects of mass deworming campaigns like one in India. ABHISEK SAHA/SOPA IMAGES/LIGHTROCKET VIA GETTY IMAGES

Even if a study isn't industry-funded, individual reviewers may have conflicts of interest. **Cochrane's policies** sometimes allow researchers who have financial ties with a company—such as grants, fees, and stocks—to participate in a meta-analysis of that company's products, provided a majority of the review authors and the lead author don't have such conflicts. Most journals' policies to prevent financial conflicts of interest are even less strict.

And Cochrane itself has been charged with bias by critics who say some reviewers have an anti-industry attitude that results in overly negative assessments of drugs and vaccines. In a **2017 review** on the effect of a new generation of antiviral drugs against hepatitis C, for instance, the authors concluded that the drugs cured the patients of the virus, but called this a "surrogate outcome"; there was no evidence that the drugs led to longer survival, they said.

The findings made many headlines, but some clinicians were outraged. The studies would have to last many years to show an effect on survival, they said, but previous trials with older drugs had clearly shown that patients who eliminate the virus live longer. "In my view, they have been too strict and they have overstated their conclusions," says Andrew Hill of the University of Liverpool in the United Kingdom. "Due to this report, patients with hepatitis C may potentially be unable to access lifesaving therapy," a group called The **Hepatitis C Coalition** wrote in *The BMJ*.

Even when no money is at stake, researchers may have an interest in the outcome of a meta-analysis—for instance because they may hope to confirm what their own studies had previously shown, or because they’ve supported certain policies.

Take the “worm wars,” over whether mass deworming campaigns among children in developing countries are clinically effective. Dozens of endemic countries have implemented mass deworming at the recommendation of the World Health Organization (WHO). But in 2015, a team led by David Taylor-Robinson of the University of Liverpool concluded in an **updated Cochrane review** that mass deworming does not improve kids’ average nutritional status, school performance, or survival; they called the belief in the positive impacts “delusional.” A year later, a team by economist Edward Miguel of the University of California, Berkeley, published **a meta-analysis** that did show clear benefits. For Miguel, who says the Liverpool study looked at the wrong things and lacked statistical power, there was more at stake than public health: He had headed several trials of mass deworming and had become a strong advocate of WHO’s policy.

Psychologist James Coyne of the University Medical Center Groningen in the Netherlands says scientists shouldn’t be involved in meta-analyses that include their own work; he has sharply **criticized** Cochrane for not taking “intellectual conflicts of interest” seriously enough. “Meta-analyses have become a tool for academics with vested interests,” he says.

When meta-analyses collide: the case of the placebo response

Some studies suggest that in recent trials of antidepressants, a growing number of patients improve just after taking a placebo, making it harder to tell whether the drug being tested is effective. Two key meta-analyses on the subject took different approaches—and reached opposing conclusions.

	FURUKAWA ET AL., LANCET PSYCHIATRY, 2016	KHAN ET AL., WORLD PSYCHIATRY, 2017
Number of studies included	 252	 85
Type of studies	Every published and unpublished randomized clinical trial done between 1978 and 2015 that the team was able to collect	Only studies reported in Food and Drug Administration reviews, for drugs approved between 1987 and 2013
Number of patients on placebo	 12 × 2000 = 24,000	 4 × 2000 = 8,000
Outcome measure	Proportion of patients who had a 50% or greater reduction of symptoms	Average decrease of symptoms, expressed as a percentage
Statistical method	Metaregression	Linear regression
Key finding	Placebo response stable since 1991; on average, 36% of patients on placebo were cured.	Average placebo response has increased by 6.4 percentage points since 2000.

Miguel says his own papers shouldn't disqualify him. "We should judge the analysis based on the quality and have the debate on scientific terms," he says. But Cochrane Library Editor-in-Chief David Tovey acknowledges the problem. Cochrane does not ban scientists from taking part in meta-analyses that include their own work, but they can't be involved in the assessment of their trials—which sometimes means having to leave the room temporarily—and the conflict has to be acknowledged in the review paper. "We recognize that this in itself is insufficient," Tovey says. "We have a new proposal coming up to make it more comprehensive. But I have to say it is jolly challenging." Finding suitable authors for a meta-analysis is hard when the people with the most expertise in an area are excluded, he explains.

Both sides in the war over media violence had a stake in the outcome. But the main reason their meta-analyses diverged was the different ways the researchers handled publication bias, the well-known phenomenon in which studies that come up empty-handed are less likely to be published.

Like other scientists who conduct meta-analyses, Ferguson and Kilburn have used several statistical methods to measure publication bias and correct for it. In one method, they plotted the outcomes of all the studies against their sample size on a graph. Without publication bias, the results would have been distributed symmetrically, and the plot would look like an inverted funnel, centered on the mean. But it didn't; the plot was lopsided. To correct for this bias, they essentially added a "missing"—and supposedly unpublished—study for each study that lacked a counterpart on the other side of the mean. With that and other corrections, the evidence that games and movies made people more aggressive evaporated, they concluded in their 2009 meta-analysis.

Bushman and Anderson took a different approach: They tried to find all unpublished studies, mainly by asking the authors of published studies whether they had failed to publish others and checking Ph.D. theses for chapters not published in scientific journals. They then included what they had collected in their meta-analysis. Applying a statistical method to show that the results of the studies were now distributed evenly around the mean reassured them that they had overcome publication bias. The apparent link between video games and aggression persisted.

The debate became heated. Ferguson accused his opponents of only collecting unpublished studies with desirable results and "overestimating and overadvertising" the effect—which Bushman and Anderson said was "a red herring." Many other scientists weighed in, as did game enthusiasts and opponents. Both sides also pushed their results outside science. Last year, Ferguson **published a book** called *Moral Combat: Why the War on Violent Video Games Is Wrong*. (The title was a play on the video game *Mortal Kombat*.) Bushman was a member of former President Barack Obama's committee on gun violence, testified before the U.S. Congress on the topic of youth violence, and has frequently been interviewed on TV.

“ **Scientists have to make several decisions and judgment calls that influence the outcome of a meta-analysis. ... Anyone who wants to manipulate has endless possibilities.** ”

Jos Kleijnen, Kleijnen Systematic Reviews

Things got even more complicated after a third researcher joined the fray in 2016: Joseph Hilgard, a psychologist at Illinois State University in Normal. Hilgard, who studies pathological aspects of gaming, including addiction, says, “I was curious: Would I be more persuaded by one or the other? So I tried to find the answer by mashing around in the data.” (Questions he raised about another paper in 2017 led to one of Bushman’s three retractions.)

With two colleagues, Hilgard reexamined Bushman’s 2010 meta-analysis, applying several novel statistical techniques to correct for publication bias, including one developed by Van Aert. Based on those results, they concluded in **a 2017 paper** in the *Psychological Bulletin* that Bushman and Anderson hadn’t managed to collect all unpublished studies, and that publication bias still played a role. After correcting for that bias, the relationship between violent games and aggression turned out to be “very small,” they said. Bushman and Anderson **reject Hilgard’s analysis** and stand by the results of their meta-analysis.

The many battles have sobered Glass, the inventor of the meta-analysis. “I have come to think of meta-analyses as a tool to convince the undecided,” he says. “To give them something useful.” Psychologist Hannah Rothstein of Baruch College in New York City, a meta-analysis consultant who collaborated with Bushman in producing the 2010 meta-analysis, says she has not lost faith in the method—but she has changed her expectations. “We used to make meta-analyses as objective as possible. Now, we try to make them as transparent as possible,” she says. “Anyone who disagrees with a certain decision will have to be able to redo it and see if that has an influence on the results.”

Ioannidis agrees. Systematic review protocols should be published up front, he says; each analytical step and every judgment call should be reported. “If this is done, one can exactly see the degrees of freedom where the deviation is creeping in, and take that into account to attach a certain credibility to the results.” In controversial cases, he adds, rival researchers should set up a meta-analysis together. Or even better, they could forget about the many studies already published and set up new ones, using standardized protocols, and then do a meta-analysis of the results, using a methodology agreed on and published in advance.

That approach was taken to settle a long-running debate over whether self-control can be depleted, just like muscles: Researchers at 23 labs around the world conducted the same standardized experiment and carried out a meta-analysis. **Published in 2016**, it showed that the effect is close to zero, an outcome now widely accepted.

That's exactly what Hilgard advocates to settle the debate on media violence and aggression. "We cannot bear the thought of another 30 years' stalemate," he wrote. Rothstein has no illusion that this would end the controversy, but it would at least move the discussion forward, she says. "Although I don't know if Ferguson and Bushman will be able to stay in a room together without killing each other."

This story was supported by the Science Fund for Investigative Reporting.

Posted in: [Scientific Community](#)

doi:10.1126/science.aav4617

Jop de Vrieze

Jop de Vrieze is a science journalist in Amsterdam.

More from News

['Journalologists' use scientific methods to study academic publishing. Is their work improving science?](#)



[NSF spells out new sexual harassment policy: Talk to us](#)



[Antarctic glacier gets new name in wake of sexual harassment finding](#)



News from *Science* has **introduced a metered paywall**. Full access to all news content is included in **AAAS membership**.

Got a tip?

How to contact the news team

Advertisement