Journal of Behavioral
Public Administration
Vol 2(2), pp. 1-10
DOI: 10.30636/jbpa.22.41

Research Article

Rates and the judgment of government performance

Oliver James*, **Gregg G. Van Ryzin**†

Abstract: Many official statistics reported to the public appear in the form of rates, such as crimes or diseases per 100,000 people, with the choice of a base number (for example per 1,000,000, per 100,000, or per 1,000) remaining largely a matter of the choices or traditions of statistical agencies. Because prior studies have shown that people tend to judge the likelihood of an event based on the numerator alone (thus exhibiting denominator neglect), we hypothesize that ratio bias influences citizens' perceptions of risks and conditions in society when interpreting real government statistics. To probe this hypothesis, we designed a pair of survey experiments in which a sample of US adults was randomly allocated to treatment groups receiving the same official statistics about violent crime (from the FBI) and infant mortality (from the CDC) but framed as rates with different base numbers (with an additional group receiving only the absolute number of events). We find some evidence of the expected ratio bias when violent crime is framed in terms of different base numbers, but the results for infant mortality were less consistent. For both violent crime and infant mortality, however, absolute numbers led to perceptions of the greatest risk and least favorable conditions, while individual rates (per person) led to perceptions of the least risk and the most favorable conditions. These findings suggest that citizens' substantive judgments about risks and conditions in society may be influenced to some extent by the framing of rates by government statistical agencies when reporting official statistics to the public.

Keywords: Ratio bias, Rates, Risk perception, Performance measurement, Accountability

Supplements: Open data, Open materials

Governments around the world regularly report official statistics in the form of rates or ratios. For example, government statistical agencies report rates for the economy (unemployment rates), public safety (crime rates), education (graduation rates), and health (rates of morbidity and mortality from various diseases or health conditions) (Horn, 1993; Hatry, 2006; Larsen & Olsen, 2019; Moynihan, 2008). Prominent examples in the United States include the Federal Bureau of Investigation (FBI) reporting on crime rates for various types of crime, and the Centers for Diseases Control and Prevention (CDC) reporting on various disease rates and health conditions. Internationally, progress on the United Nations Sustainable Development Goals is reported on measures including rates about health, environmental, and economic outcomes. And the World Bank's Open Data project (data.worldbank.org) provides access to hundreds of social, economic, education, and health indicators, many of which appear in the form of rates or ratios. These are just a few examples from the numerous statistical agencies at the local, state, national, and international levels of government around the globe that use rates and ratios routinely in reporting official statistics to the public.

The types of rates reported by governments vary across policy areas and institutions; for example, outcomes or events are sometimes reported per million, per hundred thousand, per thousand, per hundred (or percent), or per unit. Often the choice of a base depends on the frequency of occurrence of the outcome involved. For example, it would not make sense to report infant mortality as a percent (per 100) because, thankfully, in most countries infant mortality (a death during the first year of life) occurs less than once in 100 live births on average. The rate of infant mortality in the US (at the time of this study),

* University of Exeter
† Rutgers University, Newark
Address correspondence to Gregg G. Van Ryzin at (vanryzin@rutgers.edu)

for example, was 5.8 per 1,000 live births or, as more typically reported by the CDC, 582 per 100,000 live births. The choice of the base number for a rate often reflects the technical judgments and professional traditions of statistical agencies. This administrative decision may matter, however, because of evidence showing that the framing of rates can influence how information is received, processed, and used in decision making.

Indeed, psychological research has found evidence of a ratio bias in which the perceived likelihood of an event appears greater when it is presented as a ratio with a large numerator and denominator than when the same information is presented with a small numerator and denominator (Reyna, Nelson, Han, & Dieckmann, 2009). For example, Yamagishi (1997) gave participants mortality rates by varying both the percentage incidence rate and population frame (deaths per 100 or 10,000 people) within subjects. Ratings of risk were higher with a frame of 10,000 than a frame of 100, regardless of the actual percentage incidence rate. In the same manner, cancer was rated as riskier when described as killing 1,286 out of 10,000 people (12.86%) than as killing 24.14 out of 100 people (24.14%), clearly showing ratio bias in risk perception. As Pedersen (2016) notes, phenomena similar to ratio bias have variously been described as 'denominator neglect' (Okan, Garcia-Retamero, Cokely, & Maldonado, 2012), 'numerosity effect' (Reyna & Brainerd, 2008), and 'unit effect' (Pandelaere, Briers, & Lembregts, 2011). The most frequently discussed reason for ratio bias is denominator neglect (as discussed by Stone et al., 2018), which refers to the idea that the ratio is considered in its component parts of numerator and denominator, with the numerator receiving more weight. This may occur because the numerator appears first (or on top) in the expression of a ratio and because of the cognitive complexity of dividing by the denominator. Because of the resulting neglect of the denominator, equivalent ratios with larger numbers in them result in increased perceptions of risk (for example) compared to ratios with smaller numbers.

The importance of ratio bias to public policy and administration has begun to be recognized (Pedersen, 2016) and, we argue, may matter for reporting government performance and, in turn, democratic accountability. Evidence suggests that government performance reporting in general can influence citizens' perceptions of outcomes and conditions in society (James, 2011; Larsen & Olsen, 2019; Van Ryzin & Lavena, 2013). Moreover, a range of biases have been detected in how users interact with performance information (Andersen & Hjortskov, 2015; James & Van Ryzin, 2016; 2017; Olsen, 2015). Pedersen (2016) notes that numeric biases in framing policy were until recently relatively neglected in the literature on framing, which instead concentrated on framing in the sense of the wording of policy choices. In this study, we hope to contribute to the empirical literature by examining ratio bias in the reporting of rates of government performance using real examples of two important conditions in society: violent crime and infant mortality. In our experiments, a sample of US adults was randomly allocated to treatment groups receiving the same official statistics about violent crime (from the FBI) and infant mortality (from the CDC) but framed as rates with different base numbers (with one group receiving the absolute number of crimes or deaths). Both of these examples involve statistics that appear widely in reports by government and the news media, capture important societal concerns, and invoke realism and relevance to most citizens.

Given evidence from psychology of ratio bias, we propose the following main hypothesis: *reporting violent crimes and infant mortality using a larger denominator will lead citizens to perceive more risk and worse conditions in society, compared to reporting the mathematically same statistic using a smaller denominator*. That is, we expect to see ratio bias in citizens' interpretations of FBI violent crime statistics and in CDC infant mortality statistics when framed using mathematically equivalent ratios but with the outcome totals divided by differing denominators (per 1000, per 10,000, per 100,000 and per 1,000,000), with the larger denominators associated with greater perceptions of risk and lower ratings of conditions in society. We also include an individual risk ratio, for example violent crime per person or infant death per live birth, with the corresponding expectation that (as the smallest possible denominator) this will lead to the least perceived risk and the most positive ratings of conditions in society, relative to equivalent ratios with larger denominators.

We further consider the alternative condition of reporting the absolute number of violent crimes or infant deaths. This allows us to consider the effect of reporting rates compared to reporting the absolute number of outcomes or events in society. In both cases, the figures for violent crime and infant deaths are large numbers in comparison to the rates (each rate incorporates the absolute number in its calculation, of course, with a divisor larger than 1). However, the absolute number would still enable citizens to form meaningful perceptions of risk and conditions in society to the extent they can interpret the numbers in terms of their existing knowledge about the general size of the US population or number of births in the nation in a year. Of course, this is likely to be even more difficult to do then when the base is explicitly given, suggesting that citizens viewing just the absolute numbers are even more likely to disregard the (implied) denominator and thus overestimate the magnitude relative to the population. This suggests the following additional hypothesis: *reporting totals increases perceptions of risk and decreases ratings of conditions in society, compared to reporting rates.*

Finally, numeracy—the ability to correctly interpret quantitative information—has been shown to affect the consideration of statistical information of various kinds, with more numerate individuals having more skills to process such information (Peters, 2006). Indeed, previous experiments have shown that lower numeracy increases susceptibility to ratio bias (Reyna & Brainerd 2008). For this reason, we might expect that more numerate citizens would be less susceptible to ratio bias when interpreting official statistics. This leads to a third, mediation hypothesis: *ratio bias will be moderated by numeracy; specifically, more numerate citizens will be less prone to ratio bias than less numerate citizens.*

Experimental Design and Participants

To probe these hypotheses, we designed and implemented a pair of survey experiments and embedded them in an omnibus online survey of US adults conducted in February 2018. Both experiments had the same design with one presenting statistics on violent crime and the other infant mortality (see Supplement, Appendix A). We used the most recently available real US government statistics from the Federal Bureau of Investigation (FBI) Uniform Crime Reports, and the Centers for Disease Control and Prevention (CDC) National Center for Health Statistics. Thus, the statistics shown to respondents reflected actual conditions in US society at the time of the survey.

Violent crime statistics are almost always presented by the FBI as reported offenses per 100,000 inhabitants, with a nationwide rate for the US at the time of the design of the survey of 401.3 (Federal Bureau of Investigation, 2016, Table 11). Infant mortality is reported as deaths per 100,000 live births, with a nationwide rate at the time of survey design of 582.1 (National Center for Health Statistics, 2015). It should be noted that infant mortality is sometimes also reported in the form of deaths per 1,000 live births (see National Center for Health Statistics, 2016). In both experiments, we preserved the real ratio but varied the base of the ratios across experimental groups as follows: 1,000,000; 100,000; 10,000; and 1,000. We further included a group presented a unit rate of violent crime per inhabitant (.004) or infant deaths per live birth (.0058) and another group given information as the absolute number of violent crimes in the US (1,223,265) or total infant deaths in the US (23,215). Thus, in each experiment, there were 6 different ways of presenting the same official statistics, with respondents randomized to one of the six conditions. See the Appendix for actual question wording and presentation of the statistics. The order of presentation of the two experiments was randomized within subjects.

In the experiment regarding violent crime statistics, we rounded off 401.3 per 100,000 to 400 per 100,000 and extended this rounding to the other ratios as follows: 4,000 per 1,000,000; 40 per 10,000; and 4 per 1,000. We did this in part because the actual number was easy to round off without diminishing accuracy but also, more importantly, to examine the pure effect of the ratio scale and not the precision or particular configuration of the numbers in the numerator. In the infant mortality experiment, however, we did not round off the rate of 582.1 per 100,000 and extended a similar level precision to the other ratios as follows: 5,821 per 1,000,000; 58.2 per 10,000; and 5.8 per 1,000. Thus, one experiment included simple, round numbers and the other slightly more complex, precise numbers for respondents to consider. This approach was taken because the degree of precision of numbers has been shown to matter for performance reporting (Olsen, 2018) suggesting the possibility that ratio bias may differ according to amount of precision in reporting.

In each experiment, after viewing the statistical information, respondents were asked to rate perceived risk on a 0-10 scale (from "no risk at all" to "high risk"). They were also asked to rate the condition in society, on a 1-7 scale, based on the statistics shown (from "extremely bad" to "extremely good"). See Supplement, Appendix A, for the exact question wording and response formats. We also included a self-reported measure of subjective numeracy as a possible moderator (see Supplement, Appendix B), with the expectation, as discussed above, that more numerate citizens (because they are more comfortable with percentages and fractions) would exhibit less ratio bias. Although subjective scales are not the same as objective measures of numeracy (such as brief math tests), they have the advantages of taking less time and imposing less burden on survey respondents. Moreover, studies show that subjective and objective measures are strongly correlated (Zikmund-Fisher et al., 2007).

The online survey used Qualtrics software and the Research Now market research panel (administered by Qualtrics). There were n=841 valid responses from a nonprobability sample of US adults, after excluding 53 low-quality respondents who sped through the questionnaire. Quotas for region, age, gender, and race based on American Community Survey made the sample diverse and matched to the US population. The two experiments were part of the same block of questions within the survey instrument, which included other short survey experiments (in separate blocks randomized within the survey).

Analysis and Results

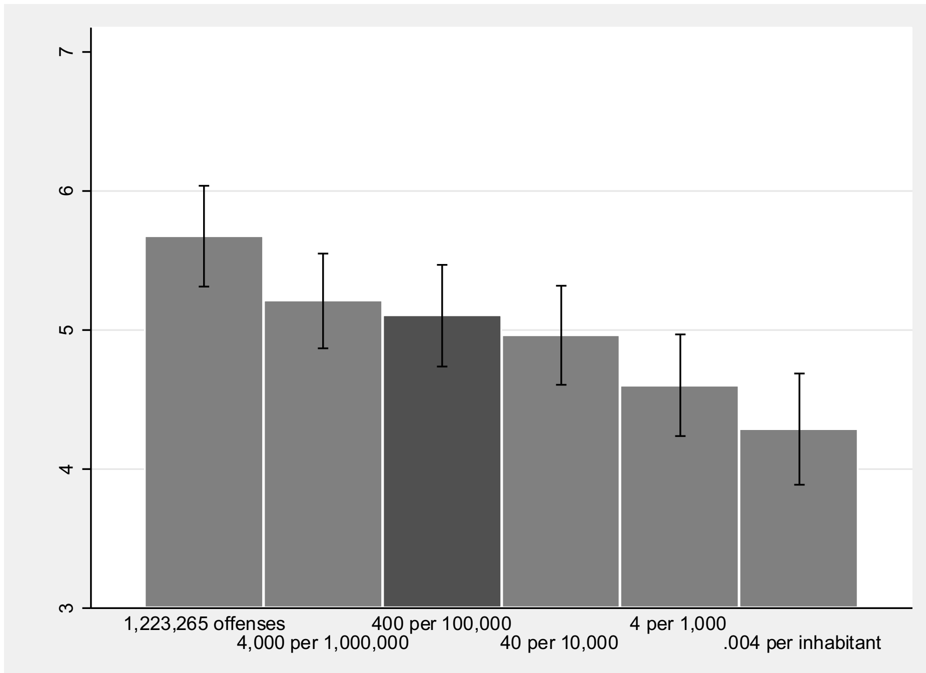
For each experiment, we begin with descriptive graphs of the means for each dependent variable across the six treatment groups, along with one-way ANOVA tests, to visualize the pattern of results and test for an overall treatment effect. We then report regressions that test the difference of treatment group means from a reference group, defined as the government's standard reporting ratio (per 100,000) used in actual public reports by the FBI and CDC. In addition to the usual unstandardized regression coefficients, we also report y-standardized coefficients as a measure of effect size. We also ran ad-hoc multiple comparisons (provide in Appendix C in the Supplement) to examine differences between any two treatment groups, largely for reference purposes.

Figure 1 shows the descriptive results for perceived risk of violent crime across treatment groups in the FBI experiment. A clear pattern can be seen of lower perceived risk when respondents viewed smaller-number ratios (one-way ANOVA $F = 4.80$, $p < .001$, see Table C1 in the Supplement). The individual risk, .004 violent crimes per inhabitant, is judged to be the lowest risk. In contrast, the total number of violent crimes, 1,223,265, is judged to be the highest risk. Figure 2 shows the results for the rating of conditions in society across treatment groups, which is our second dependent variable, and here the pattern is the inverse: smaller-number ratios are associated with higher (better) ratings of conditions (one-way ANOVA, $F = 12.12$, $p < .001$, see Table C2 in the Supplement). The individual risk ratio, .004 violent crimes per inhabitant, leads to the most favorable rating of conditions in society. The total number of violent offenses, 1,223,265, leads to the least favorable rating of conditions of society.

Table 1 presents the linear regressions for the presentation of FBI statistics on violent crimes, with separate linear regressions for perceived risk and condition ratings. The experimental groups are represented as dummy variables with the excluded group, as mentioned previously, the FBI's standard ratio for reporting crime (per 100,000 inhabitants). Thus, the coefficients in Table 1 represent contrasts with this standard ratio. Tests of comparisons between all pairs of groups are also provided in Appendix C in the Supplement for reference. As Table 1 shows, the level of perceived risk when respondents are exposed to the standard ratio of 400 reported offenses per 100,000 inhabitants (the constant) is 5.1 on the 1-10 perceived risk scale. Presenting 4,000 reported offenses per 1,000,000 inhabitants leads to a slightly higher level of perceived risk but not a significant difference from the standard ratio. Correspondingly, presenting 40 reported offenses per 10,000 inhabitants, and moreover 4 reported offenses per 1,000 inhabitants, results in a lower level of perceived risk but not significantly different from the standard ratio. However, offenses per inhabitant (.004) does lead to a significantly lower level of perceived risk (-0.815 scale points, or -0.308 SDs) compared to the constant, while total reported violent offenses (1,223,265) leads to a significantly higher perceived risk (+0.571 scale-points, or +0.216 SDs) compared to the constant. Thus, although Figure 1 corresponds to what ratio bias theory predicts, only the extremes of the individual risk ratio, on the one end, and the total violent crimes, on the other end, lead to judgments of risk that differ significantly from the presentation of the FBI's standard ratio of 400 violent crimes per 100,000 inhabitants.

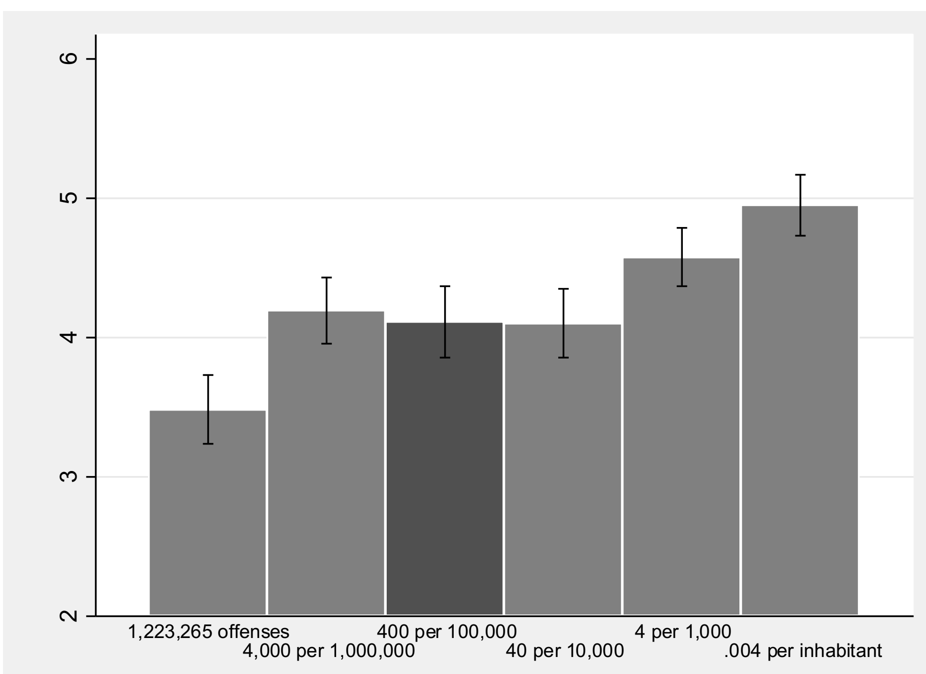
The results for the rating of crime conditions in the US are shown in the rightmost two columns in Table 1. The FBI standard reported rate (400 per 100,000) is again the excluded group and thus captured in the constant, which has a mean of 4.1 on the 1-7 conditions rating scale. Presenting violent crimes as either 4,000 per 1,000,000 or as 40 per 10,000 does not lead to statistically significant differences in the ratings of conditions in society. However, respondents shown the crime statistics in the form of 4 reported offenses per 1,000 inhabitants rate conditions in society as better (+0.464 scale-points, or +.265

Figure 1
Perceived Risk, FBI Reports on Violent Crime



Note: Error bars indicate 90 percent confidence intervals.

Figure 2
Condition Rating, FBI Reports on Violent Crime



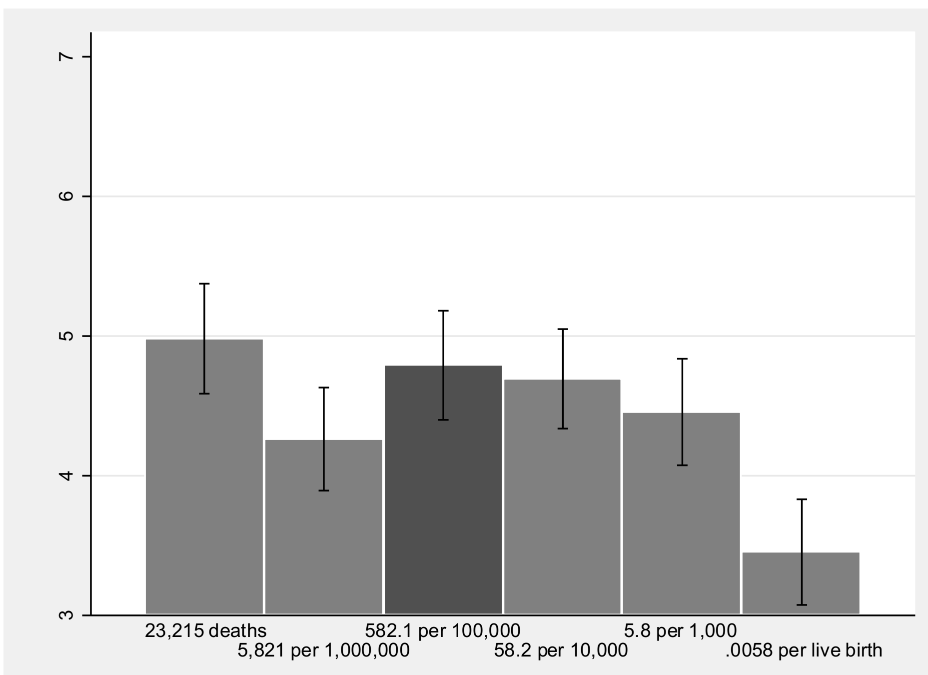
Note: Error bars indicate 90 percent confidence intervals.

Table 1
Regression Analysis of FBI Reports on Violent Crime

	Perceived risk (1-10 scale)			Condition rating (1-7 scale)		
Reported offenses	Unstand. Coeff.	Y-std. Coeff.	p-value	Unstand. Coeff.	Y-std. Coeff.	p-value
Reported offenses: 1,223,265	0.571	0.216	0.070	-0.628	-0.359	0.002
4,000 per 1,000,000 inhabitants	0.106	0.040	0.737	0.083	0.047	0.686
400 per 100,000 inhabitants (constant)	5.103	0.049		4.110	-0.071	
40 per 10,000 inhabitants	-0.140	-0.053	0.659	-0.009	-0.005	0.966
4 per 1,000 inhabitants	-0.503	-0.190	0.110	0.464	0.265	0.023
.004 per inhabitant	-0.815	-0.308	0.010	0.840	0.480	0.000
	R^2=0.028 (n=831)			R^2=0.067 (n=839)		

Note: Both unstandardized and y-standardized coefficients (as a measure of effect size) shown.

Figure 3
Perceived Risk, CDC Reports on Infant Mortality



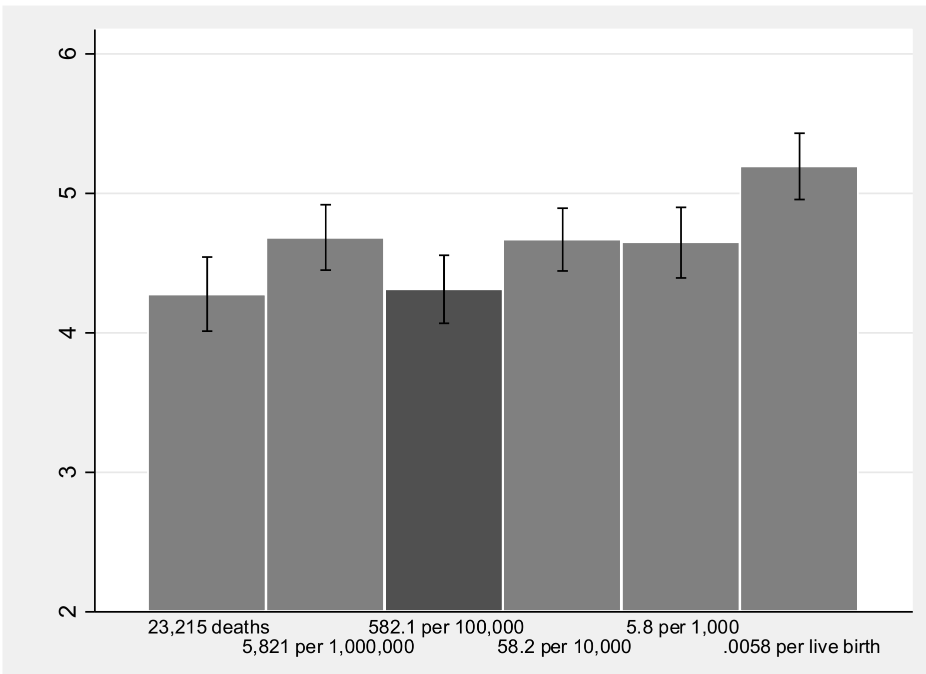
Note: Error bars indicate 90 percent confidence intervals.

SDs). Presentation of rates as offenses per inhabitant (.004) results in an even higher rating (+0.84 scale-points, or +.48 SDs) compared to the constant, a difference that is statistically significant. The total number of reported violent offenses (1,223,265) has the lowest condition rating (-0.63 scale-points, or -.359 SDs) compared to the constant, a difference that is also statistically significant.

Turning to the experiment involving CDC statistics on infant mortality, Figure 3 shows the means of perceived risk across treatment groups. The pattern is not as clear (in terms of decreasing perception of risk with smaller number ratios) as it was for the FBI statistics, but still there is a significant overall difference in risk perceptions across groups (ANOVA $F = 5.6$, $p < .001$, see Table C3 in the Supplement).

Mortality per live birth (.0058) results in the lowest risk perception and the total number of infant deaths (23,215) results in the highest risk perception. Perceived risk diminishes slightly when a smaller-number ratio is presented, except for the group shown 5,821 infant deaths per 1,000,000 live births, which defies the pattern. Figure 4 presents the mean rating of conditions in society across treatment groups, and again the pattern is mixed and not very clearly one of improving perceptions of conditions with smaller number ratios. Still, the overall difference in means across treatment groups is statistically significant (ANOVA F = 5.01, p < .001, see Table C4 in the Supplement). Total infant deaths (23,215) leads to the lowest rating of conditions, while the ratio per live birth (.0058) leads to the highest rating of conditions. But the means of the remaining treatment groups appear fairly similar.

Figure 4
Condition rating, CDC reports on infant mortality



Note: Error bars indicate 90 percent confidence intervals.

Table 2
Regression analysis of CDC reports on infant mortality

	Perceived risk (1-10 scale)			**Condition rating (1-7 scale)**		
Infant deaths	Unstand. Coeff.	Y-std. Coeff.	p-value	Unstand. Coeff.	Y-std. Coeff.	p-value
23,215	0.186	0.068	0.570	-0.038	-0.021	0.857
5,821 per 1,000,000 live births	-0.529	-0.193	0.105	0.369	0.209	0.078
582.1 per 100,000 live births (constant)	4.791	0.128		4.312	-0.180	
58.2 per 10,000 live births	-0.097	-0.035	0.763	0.355	0.201	0.087
5.8 per 1,000 live births	-0.337	-0.123	0.297	0.333	0.189	0.109
.0058 per live birth	-1.339	-0.488	0.000	0.881	0.499	0.000
	R^2=0.033 (n=832)			R^2=0.029 (n=839)		

Note: Both unstandardized and y-standardized coefficients (as a measure of effect size) shown.

Table 2 presents the regression analysis for the presentation of CDC statistics on infant mortality. In both models reported in Table 2, the ratio of 582.1 per 100,000 live births is used as the excluded category and thus the coefficients represent contrasts with the mean for this group. With respect to perceived risk, only infant deaths per live birth (.0058) leads to risk perceptions that are significantly different (-1.339 scale-points, or -.488 SDs) from the constant. With respect to the rating of conditions in US society, again only the presentation of infant deaths per live birth (.0058) leads to clearly significantly different condition ratings (+0.881 scale-points, or +.499 SDs) compared to the constant. The remaining contrasts are only borderline significant.

Finally, we examined the potential moderating role of subjective numeracy using ANCOVA, with the treatment groups as the factor variable and a scale of subjective numeracy as the covariate (see Appendix D). The subjective numeracy scale is internally consistent (alpha = .91) with a mean of 5.0 (SD = 1.7) on the 1-7 self-reported scale averaged across the four items. For FBI violent crime statistics, the ANCOVA results show that subjective numeracy has a significant main effect on risk perceptions ($F=24.98$, $p<.001$) and a significant main effect on crime conditions ($F=63.92$, $p<.001$). However, there was no significant interaction effect for either risk perception ($F=0.55$, $p=.742$) or the rating of conditions in society ($F=1.28$, $p=269$). (See Tables D1 and D2 in the Supplement.) Thus, subjective numeracy does not moderate the bias in judgments of rates in the experiment with FBI violent crime statistics. For the CDC infant mortality statistics, subjective numeracy has no main effect on risk perception ($F=0.77$, $p=.380$) but does have a significant main effect on the rating of conditions in society ($F=17.51$, $p<.001$). But there were no significant interaction effects for either risk perception ($F=0.89$, $p=.484$) or the rating of conditions in society ($F=0.72$, $p=.608$). (See Tables D3 and D4 in the Supplement.) Thus, again, subjective numeracy does not moderate the bias in judgments of rates in the experiment with CDC infant mortality statistics.

Discussion

Evidence from our experiments using a nationwide sample of US adults shows that citizens' judgments of rates are somewhat susceptible to ratio bias, although the pattern is mixed and the magnitude of bias is small. Specifically, we found that presenting a unit rate (violent crimes per inhabitant or infant mortality per live birth) resulted in clearly lower perceived risk and more favorable evaluations of conditions in society, compared with the FBI and CDC's standard reporting ratio of per 100,000. Evidence for ratio bias when other rates are compared with this standard ratio was more mixed and rather small in magnitude. These results are partially consistent with what has been called numerosity bias or nominator neglect and extends evidence of this ratio bias found in other domains. The findings are also consistent with evidence of ratio bias in support for public policies (Pedersen, 2015). We found no evidence that people's numeracy reduces bias in their perceptions of information framed according to different rates, suggesting that ratio bias is a fairly general phenomenon and may even affect experts. Indeed, studies of ratio bias involving public managers and other professionals would an interesting line of future research.

Reporting absolute frequencies (the total number of violent crimes or infant deaths) resulted in the highest perceived risk and lowest ratings of conditions in society. For the FBI violent crime statistics, reporting total crimes produced a higher perception of risk and a lower rating of conditions in society. However, the difference in perceptions of crime risk between the absolute number and the FBI standard rate per 100,000 was only borderline significant statistically. For the CDC statistics on infant mortality, the lack of significant differences between the total and the standard CDC rate per 100,000 may be related to the smaller overall number of infant deaths (23,215 compared to 1,223,265 violent crimes), the greater salience of crime perhaps, or possibly to the fact that the infant mortality numbers were not as neatly rounded off as the violent crime numbers. The extent to which simpler, rounded-off numbers versus more precise, complex numbers influences ratio bias is another interesting topic for potential future investigation.

Our findings suggest other, broader avenues for further research on ratio bias in reporting government performance. Presentational forms that de-bias perceptions could be examined. For example, the use of physical visual displays depicting the numbers contained in ratios has been found to reduce ratio bias (Stone et al., 2018). The role of bias in other ways of framing rates in reporting could also be examined, for example in reports of the kind that are often made in comparing schools, hospitals or local government units. Such comparisons often include performance over time for the same unit, comparisons between units, and comparisons both over time and across units. On this basis, a difference in reported comparative performance may be seen as a bigger and more persuasive if the numbers used in the rate are larger, even for logically equivalent rates of comparison across units or over time. This speculation would be interesting to test empirically.

Importantly, our evidence suggests the subtle but potentially important influence the framing of rates by public managers or politicians can have on citizens' judgments of risks and conditions in society. This finding opens up an area of research about whether the framing of rates may be actively pursued by public officials, similar to evidence showing that politicians sometimes exploit cognitive biases in the way citizens process rates of taxation (Krishna & Slemrod, 2003; Olsen, 2013). This possibility further adds to the list of difficulties in government of reporting crime statistics to citizens (Larsen & Olsen, 2019; Mosher, Miethe, & Hart, 2010). Indeed, our results clearly show that the FBI reporting of violent crime per 100,000 (which is also the rate typically reported in the media, Lowry et al., 2003) tends to bring about a higher assessment of risk and worse perception of conditions compared to the alternative of reporting, particularly the violent crime rate per inhabitant. Thus, the framing of rates could be explored as a potential factor behind the salience of crime as a perceived problem in the US—and the related perception that government is not doing enough to address the problem (Hetherington, 2005; Hetherington & Rudolph, 2008). Reporting crime per 1,000 people, or reporting both metrics, could be informative in giving citizens different perspectives. However, the legitimacy of intentionally framing rates to alter citizens' judgments is debatable. Attempts to manipulate rates in this way could undermine citizens' trust in official statistics, adding to the credibility problem of government performance (Van Ryzin & Lavena, 2013; James & Van Ryzin, 2017). Relying on official statistical agencies, such as inspectorates and audit bodies, to decide the best framing of rates to report to the public certainly helps reduce the potential for political gaming and manipulation. But even politically neutral statistical agencies need to be aware of how the rates they use may still influence the perceptions and judgments of the public.

References

Alonso, A., & Fernandez-Berrocal, P. (2003). Irrational decisions: attending to numbers rather than ratios. *Personality and Individual Differences*, *35*(7), 1537–1547.

Andersen, S.C. & Hjortskov, M. (2015). Cognitive biases in performance evaluations. *Journal of Public Administration Research and Theory*, *26*(4), 647-662.

Bonner, C., & Newell, B. R. (2008). How to make a risk seem riskier: The ratio bias versus construal level theory. *Judgment and Decision Making*, *3*(5), 411–416.

Federal Bureau of Investigation (2017). *Crime in the United States*, 2016. US Department of Justice.

Gilens M. (2001). Political ignorance and collective policy preferences. *American Political Science Review*, *95*(2): 379–396.

Hatry, Harry P. (2006). *Performance measurement: Getting results*, 2nd Edition. Washington DC: The Urban Institute Press.

Hetherington, Marc J. (2005). *Why trust matters: Declining political trust and the demise of American liberalism*. Princeton, N.J.; Woodstock: Princeton University Press.

Hetherington, M. J., & Rudolph, T. J. (2008). Priming, performance, and the dynamics of political trust. *The Journal of Politics*, *70*(2), 498-512.

Horn, R. V., 1993. *Statistical indicators for the economic and social sciences*. Cambridge, Cambridge University Press.

James, O. (2011). Performance measures and democracy: Information effects on citizens in field and laboratory experiments. *Journal of Public Administration Research and Theory, 21*(3), 399-418.

James, O., & Van Ryzin, G. G. (2016). Motivated reasoning about public performance: An experimental study of how citizens judge the affordable care act. *Journal of Public Administration Research and Theory*, *27*(1), 197-209.

James, O. & Van Ryzin, G.G. (2017). Incredibly good performance: An experimental study of source and level effects on the credibility of government. *The American Review of Public Administration, 47*(1), 23-35.

James, O., Jilke, S., & Van Ryzin, G.G. (2017). *Experiments in public management research: Challenges and contributions*. Cambridge, Cambridge University Press.

Krishna, A. & Slemrod, J. (2003). Behavioral public finance: Tax design as price presentation. *International Tax and Public Finance, 10*(2), 189–203.

Larsen, M. V., & Olsen, A. L. (2019). Reducing Bias in Citizens' Perception of Crime Rates: Evidence from a Field Experiment on Burglary Prevalence. *Journal of Politics*.

Lawrence, E. D. & Sides, J. (2014). The consequences of political innumeracy. *Research & Politics, 1*(2), 1–8.

Lowry, D. T., Nio, T. C. J., & Leitner, D. W. (2003). Setting the public fear agenda: A longitudinal analysis of network TV crime reporting, public perceptions of crime, and FBI crime statistics. *Journal of Communication, 53*(1), 61-73.

Mérola V. & Hitt M. P. (2015). Numeracy and the persuasive effect of policy information and party cues. *Public Opinion Quarterly (Forthcoming).* doi:10.1093/poq/nfv051

Mosher, C. J., Miethe, T. D., & Hart, T. C. (2010). *The Mismeasure of Crime*. Thousand Oaks, CA: Sage.

Moynihan, D. P. (2008). *The dynamics of performance management: Constructing information and reform*. Georgetown University Press.

National Center for Health Statistics. (2015). *Mortality in the United States in 2015*.

Okan, Y., Garcia-Retamero, R., Cokely, E. T., & Maldonado, A. (2012). Individual differences in graph literacy: Overcoming denominator neglect in risk comprehension. *Journal of Behavioral Decision Making*, *25*, 390–401.

Olsen, A. L. (2013). The politics of digits: Evidence of odd taxation. *Public Choice*, *154*(1-2), 59–73.

Olsen, A.L. (2015). Citizen (dis)satisfaction: An equivalence framing study. *Public Administration Review, 75*(3), 469–78.

Olsen, A. L. (2018). Precise performance: Do citizens rely on numerical precision as a cue of confidence? *Journal of Behavioral Public Administration*, *1*(1), 1-10.

Pandelaere, M., Briers, B., & Lembregts, C. (2011). How to make a 29% increase look bigger: The unit effect in option comparisons. *Journal of Consumer Research*, *38*(2), 308–322.

Peters E. (2006). Beyond comprehension: The role of numeracy in judgments and decisions. *Current Directions in Psychological Science, 21*(1): 31–35.

Reyna, V. F. & Brainerd, C. J. (2008). Numeracy, ratio bias, and denominator neglect in judgments of risk and probability. *Learning and Individual Differences, 18*(1), 89–107.

Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychological Bulletin, 135*(6), 943–73.

Stone, E. R., Parker, A. M., & Townsend, L. D. (2018). Distinguishing the ratio bias from unsystematic error: Situation and individual‐difference effects. *Journal of Behavioral Decision Making*, *31*(4), 587-601.

Yamigishi, K. (1997). When a 12.86% mortality is more dangerous than 24.14%: Implications for risk communication. *Applied Cognitive Psychology, 11*(6), 495–506.

Van Ryzin, G. G. & Lavena, C. F. (2013). The credibility of government performance reporting. *Public Performance & Management Review, 37*(1), 87-103.

Zikmund-Fisher, B. J., Smith, D. M., Ubel, P. A., & Fagerlin, A. (2007). Validation of the subjective numeracy scale: Effects of low numeracy on comprehension of risk communications and utility elicitations. *Medical Decision Making, 27*(5), 663-671.