

Playing Some Video Games but Not Others Is Related to Cognitive Abilities: A Critique of Unsworth et al. (2015)

C. Shawn Green¹, Florian Kattner¹, Adam Eichenbaum², Benoit Bediou³, Deanne M. Adams⁴, Richard E. Mayer⁵, and Daphne Bavelier^{3,6}

¹Department of Psychology, University of Wisconsin–Madison; ²Helen Wills Neuroscience Institute, University of California, Berkeley; ³Faculté de Psychologie et Sciences de l'Éducation, Université de Genève; ⁴Department of Psychology, University of Notre Dame; ⁵Department of Psychological and Brain Sciences, University of California, Santa Barbara; and ⁶Department of Brain and Cognitive Sciences, University of Rochester

Received 8/13/15; Revision accepted 3/26/16

In a recent report, Unsworth and his colleagues (2015) suggested that no relation exists between playing video games and cognitive skills. In this Commentary, we briefly summarize the substantial issues with their study—in particular, with the statistical approach and overall research methodology—that together cast serious doubt on their conclusions.

Interpreting Multiple Independent Linear Correlations Is Problematic

Unsworth et al. examined the relation between playing video games of various genres and cognitive abilities by performing more than 150 simple correlations. Specifically, each individual analysis modeled participants' scores on a given cognitive task as a simple linear function of the time the participants spent playing games of a single genre over the previous year, without controlling for the participants' experience with other genres or experience prior to the previous year. Although performing more than 150 independent simple correlations lacks statistical rigor in its own right, the approach is particularly problematic in the case of this specific data set for at least three key reasons (for a simulation illustrating the problems, see Simulation Study in the Supplemental Material available online).

First, under such an approach, a positive relation between experience with a given genre (e.g., real-time strategy games) and performance on a cognitive task will necessarily act as evidence *against* the relation between experience with any other genre (e.g., first-person shooter games) and performance on the same task. As an

analogy, consider the effect of playing different types of sports on aerobic health. There are many different sports, such as basketball, soccer, running, and swimming, that would each be expected to be positively related to aerobic health. However, under the analysis approach used by Unsworth et al., high aerobic health in avid basketball players, soccer players, and runners would actually serve as evidence against there being a positive relation between swimming and aerobic health. This is because the relation between the amount of experience with each single activity and the outcome would be considered without controlling for the other types of experience. Thus, when hours of swimming per week are plotted against aerobic health, the avid basketball players, soccer players, and runners would all be coded as having zero hours of swimming. This, in turn, would greatly reduce the ability to observe any positive effect of swimming on aerobic health.

Rather than pitting multiple sports that all involve aerobic activity against one another, Unsworth et al. pitted multiple video-game genres that all involve common mechanics against one another. For example, first-person shooter, action, and role-playing games involve common action and shooting mechanics (whether first-person or third-person).

A second issue with Unsworth et al.'s use of multiple independent linear correlations is that they did not take

Corresponding Author:

C. Shawn Green, Department of Psychology, University of Wisconsin–Madison, 1202 W. Johnson St., Madison, WI 53706
E-mail: cshawn.green@wisc.edu

into account gaming experience before the previous year. Returning to the sports analogy, we note that such an approach assumes that an individual who avidly played both basketball and soccer during high school but then did not play at all during his or her first year in college should be in all respects identical to an individual who was sedentary (i.e., zero hours of aerobic activity) his or her entire life. Such an assumption is particularly troubling in the case of gaming, as there are data showing that the effects of just a few hours of action-video-game play can last for months to years (Feng, Spence, & Pratt, 2007; Li, Polat, & Bavelier, 2007).

A third issue is that these analyses assumed a linear relation between the hours of game play per week and cognitive outcomes. However, the relation between practice-induced skill improvement and experience or time is well known to be highly nonlinear (Doshier & Lu, 2007; Heathcote, Brown, & Mewhort, 2000; Newell & Rosenbloom, 1981). This assumption of a linear form is equivalent to assuming that the difference in aerobic health between completely sedentary individuals and individuals who engage in aerobic activity 5 hr a week should be exactly the same as the difference in aerobic health between individuals who engage in aerobic activity 15 hr a week and those who engage in such activity 20 hr a week. The relatively vast literature on learning provides no support for the assumption of a linear relation between practice and outcomes.

The Video-Game-Experience Questionnaires Were Not Designed to Provide Detailed Weekly Estimates of Game Play

Although the issues we have outlined in the previous section could potentially be resolved via more appropriate statistical techniques, the data that Unsworth et al. used, unfortunately, came from questionnaires that were not designed to support any such analyses. First, it is worth noting that the questionnaire used in Experiment 2 was of Unsworth et al.'s own design, rather than a questionnaire that had been used previously in the literature. The data gathered using this questionnaire resulted in more than 45% of the population being classified as non-video-game players; this value is at least 3 times larger than what has been found with any other questionnaire used in the literature (including the questionnaire that was developed by Green and Bavelier's research group and that Unsworth et al. used in Experiment 1; we provide this questionnaire in our Supplemental Material—see Video-Game-Expertise Classification Scheme). Such a discrepancy already casts serious doubt about the validity of Experiment 2.

However, even the more standard questionnaire used in Experiment 1 was not designed to support the analyses conducted by Unsworth et al. Indeed, across domains (e.g., alcohol use) that utilize questionnaires with a similar format, it is accepted that although such quantity-frequency (QF) measures can dependably categorize extremes of behavior, their reliability is often too poor to measure finer gradations of behavior (Room, 1990; Shakeshaft, Bowman, & Sanson-Fisher, 1999). Thus, it is not surprising that Unsworth et al. replicated the advantage in attentional skills that has been consistently seen in the literature comparing avid players of first- and third-person shooter games with nongamers (i.e., when using an extreme-groups analysis, as described in Video-Game-Expertise Classification Scheme in the Supplemental Material), but failed to detect subtler relations.

The general unsuitability of using this type of measure in continuous analyses is only magnified by the fact that increasing the number of subdivisions of behavior in QF measures results in even more unreliable estimates. In particular, it leads to large overestimations of behavior (Sobell & Sobell, 2003), an effect that we recently quantified in the case of video-game-usage measures as well. As part of a large survey administered in an introductory-psychology course, we included a questionnaire similar to the one used by Unsworth et al. in their Experiment 1. A total of 824 students filled out this questionnaire, which included questions about their experience playing games of various genres. On a separate occasion, the same students were asked to report the total number of hours they spent each week playing video games (i.e., all genres lumped together). As shown in Figure 1a, the sum of the number of hours reported in the genre-by-genre questionnaire exceeded (and in some cases greatly exceeded, given the right skew of the distribution) the number of hours reported in the total-hours questionnaire. More critically, this discrepancy increased as a function of the number of distinct genres individuals reported engaging with. That is, individuals who reported playing games of only one genre gave more consistent estimates than those who reported playing games of multiple genres. For instance, individuals who played games of more than four distinct genres reported playing between 10 and 25 more hours a week when asked to estimate their video-game play genre by genre than when asked about their overall weekly game play. Note that individuals who played games of four or more distinct genres made up a significant portion of the total population (Fig. 1b), which means that the issue of unreliability of the hourly estimates for the genres cannot be mitigated simply by using larger sample sizes (i.e., the bias is not constant). Thus, even if Unsworth et al. were to utilize a more sophisticated form of analysis, the questionnaire data simply would not provide the reliability necessary to detect the relations they sought to model.

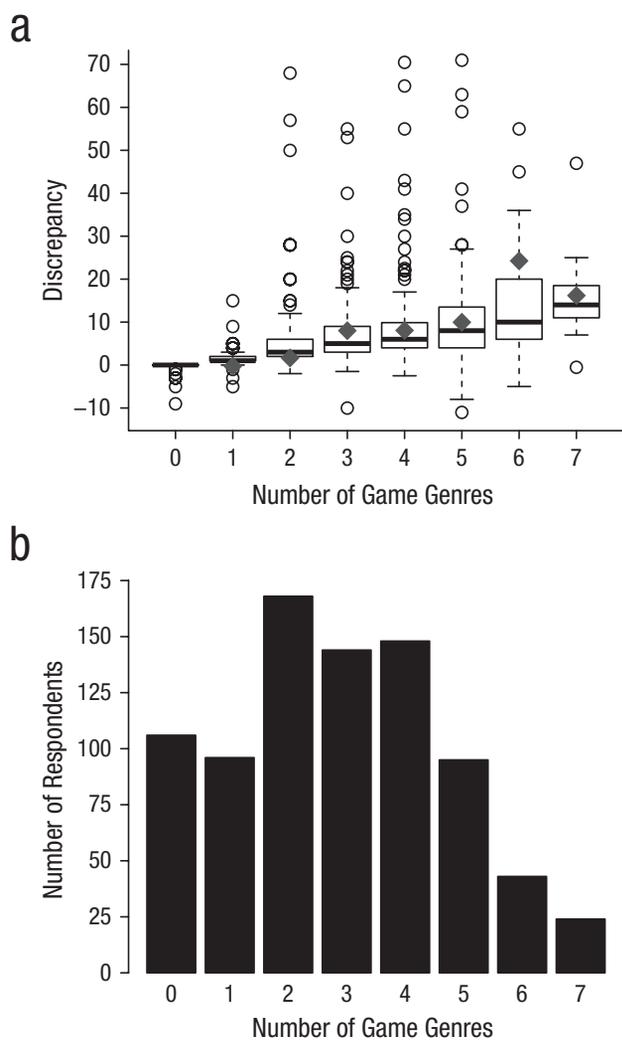


Fig. 1. Reliability of the video-game-experience questionnaires. The box-and-whiskers plot in (a) shows the discrepancy between the sum of the number of hours of video-game play respondents reported in the genre-by-genre questionnaire and the number of hours of play they reported in the total-hours questionnaire (genre-by-genre estimate minus total-hours estimate) as a function of the number of distinct genres they reported engaging with. The lower and upper edges of the boxes represent the first and third quartiles, respectively. The lines in the center of the boxes represent the medians, and the gray diamonds represent the arithmetic means. The upper and lower ends of the whiskers represent the values $1.5 \times$ the interquartile range above and below the third and first quartile, respectively, and the circles represent outliers. The graph in (b) shows the distribution of respondents according to the number of genres of video games they reported playing.

Conclusions

Although the statistical and methodological shortcomings of Unsworth et al.'s study cast significant doubt on the authors' conclusions, there remain some positive aspects to the study as well. For example, there is obvious virtue in moving toward larger sample sizes (e.g., greater reliability) and larger task batteries (e.g., ability to perform

latent-variable analyses), as Unsworth et al. did. However, in order to truly address the questions of greatest interest to the field, these steps must be taken in tandem with other changes, such as utilizing methods that more accurately measure experience (e.g., daily diaries or external monitoring) and developing better methods to describe and classify games (as separating games that have common mechanics but nominally belong to different genres will not be productive).

Action Editor

John Jonides served as action editor for this article.

Author Contributions

C. S. Green developed the study concept. C. S. Green and A. Eichenbaum developed the study design. Data collection was performed by A. Eichenbaum. A. Eichenbaum and F. Kattner performed the data analysis under the supervision of C. S. Green. C. S. Green drafted the manuscript, and all the authors provided critical revisions. All the authors approved the final version of the manuscript for submission.

Declaration of Conflicting Interests

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

Supplemental Material

Additional supporting information can be found at <http://journals.sagepub.com/doi/suppl/10.1177/0956797616644837>

References

- Dosher, B. A., & Lu, Z. (2007). The functional form of performance improvements in perceptual learning: Learning rates and transfer. *Psychological Science, 18*, 531–539.
- Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science, 18*, 850–855.
- Heathcote, A., Brown, S., & Mewhort, D. J. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin & Review, 7*, 185–207.
- Li, R., Polat, U., & Bavelier, D. (2007, May). *Action video game playing alters early visual processing*. Paper presented at the annual meeting of the Visual Sciences Society, Sarasota, FL.
- Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1–55). Hillsdale, NJ: Erlbaum.
- Room, R. (1990). Measuring alcohol consumption in the United States: Methods and rationales. In L. T. Kozlowski, H. M. Annis, H. D. Cappell, F. B. Glaser, M. S. Goodstadt, Y. Israel, H. Kalant, E. M. Sellers, & E. R. Vingilis (Eds.), *Research advances in alcohol and drug problems: Vol. 10* (pp. 39–80). New York, NY: Plenum Press.
- Shakeshaft, A. P., Bowman, J. A., & Sanson-Fisher, R. W. (1999). A comparison of two retrospective measures of weekly

- alcohol consumption: Diary and quantity/frequency index. *Alcohol and Alcoholism*, 34, 636–645. doi:10.1093/alcalc/34.4.636
- Sobell, L. C., & Sobell, M. B. (2003). Alcohol consumption measures. In J. P. Allen & V. B. Wilson (Eds.), *Assessing alcohol problems: A guide for clinicians and researchers* (2nd ed.; NIH Publication No. 03-3745, pp. 75–99). Bethesda, MD: U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health, National Institute on Alcohol Abuse and Alcoholism.
- Unsworth, N., Redick, T. S., McMillan, B. D., Hambrick, D. Z., Kane, M. J., & Engle, R. W. (2015). Is playing video games related to cognitive abilities? *Psychological Science*, 26, 759–774. doi:10.1177/0956797615570367