

Can Item Effects Explain Away the Evidence for Unconscious Sound Symbolism? An Adversarial Commentary on Heyman, Maerten, Vankrunkelsven, Voorspoels, and Moors (2019)



Hugh Rabagliati¹, Pieter Moors², and Tom Heyman³

¹School of Philosophy, Psychology & Language Sciences, The University of Edinburgh; ²Laboratory of Experimental Psychology, Department of Brain and Cognition, KU Leuven; and ³Department of Methodology and Statistics, Leiden University

Received 10/28/19; Revision accepted 7/2/20

Sound symbolism refers to the intuition that a word's sound should match the characteristics of its referent (e.g., *kiki* should label something spiky) and its prevalence and systematicity should provide compelling evidence for an intuitive mapping between linguistic form and meaning. Striking recent work (Hung, Styles, & Hsieh, 2017) suggests that these mappings may have an unconscious basis and that participants may be able to compute the fit between a word's sound and an object's shape when both are masked from awareness. This surprising finding, replicated in the preregistered report by Heyman, Maerten, Vankrunkelsven, Voorspoels, and Moors (2019), has potentially far-reaching implications for the role of awareness in language processing (Hassin, 2013; Rabagliati, Robertson, & Carmel, 2018).

However, there are significant concerns about the generalizability of those findings and whether they in fact provide evidence for the unconscious computation of sound symbolism. Here, we present an "adversarial Commentary," in which we highlight those concerns and discuss how they might be resolved. Our adversarial Commentary borrows from the idea of adversarial collaborations (Mellers, Hertwig, & Kahneman, 2001), in which scientists, despite holding conflicting theoretical views, work together to design experiments that will be maximally informative tests of their theories. In the case of our Commentary, we synthesize critiques of the original work, by Rabagliati, with reasoned counterarguments, by Heyman and Moors, in order to reach a better understanding and convey a clearer message to readers. Importantly, where traditional Commentaries and Replies may talk past one another, our adversarial

Commentary presents an integration of our views. As a result, our Commentary offers an agreed path forward for understanding unconscious sound symbolism, with broader implications for methodological standards in researching unconscious cognition.

The prior articles on unconscious sound symbolism used a technique called *breaking continuous flash suppression*, which relies on binocular rivalry. One eye is shown a rapidly changing pattern that dominates awareness and can mask the stimulus that is shown to the other eye, which in this paradigm is either a puffy or a spiky shape with either the words *kiki* or *bubu* printed inside. When the pronunciation of the word mismatched the shape of the image, both the original study (Hung et al., 2017) and the replication study (Heyman et al., 2019) found that stimuli were suppressed from awareness for longer (i.e., breakthrough times were longer for incongruent stimuli).

Breakthrough from continuous flash suppression has been used to make a number of strong claims about what can be processed without awareness, from facial emotions (Yang, Zald, & Blake, 2007) to sentence meanings (Sklar et al., 2012), but not every claim has generalized. For instance, Rabagliati et al. (2018) consistently failed to replicate findings that the meanings of words and phrases affected breakthrough but did find that

Corresponding Author:

Hugh Rabagliati, The University of Edinburgh, School of Philosophy, Psychology & Language Sciences, 7 George Square, Edinburgh, EH8 9JZ, United Kingdom
E-mail: hugh.rabagliati@ed.ac.uk

Psychological Science
2020, Vol. 31(9) 1200–1204
© The Author(s) 2020
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0956797620949461
www.psychologicalscience.org/PS

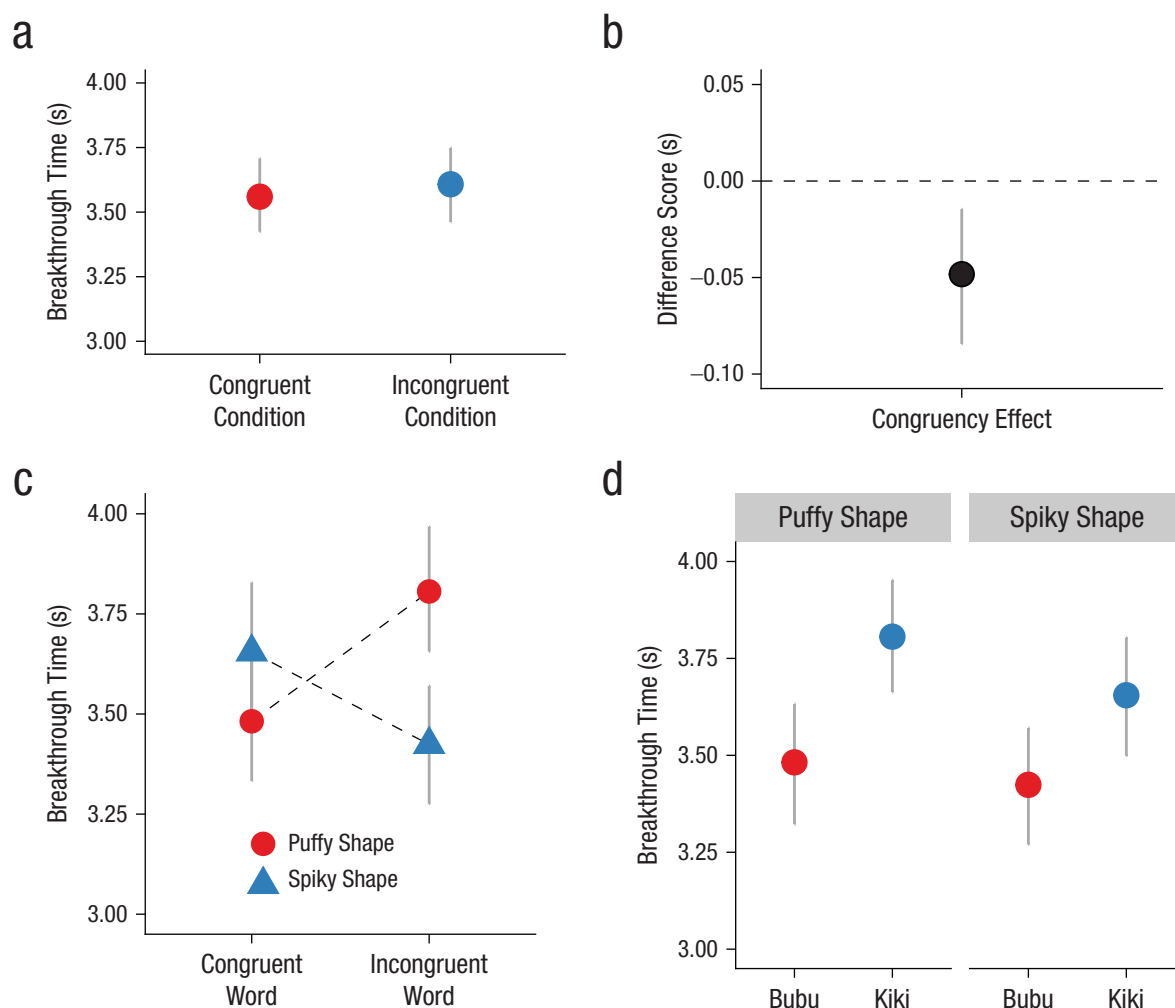



Fig. 1. Results from Heyman, Maerten, Vankrunkelsven, Voorspoels, and Moors (2019; top row) and the present analyses (bottom row). Omnibus breakthrough times in Heyman et al.'s study (a) are shown as a function of shape–word congruency. The breakthrough difference score (b) was calculated by subtracting mean incongruent breakthrough times from mean congruent breakthrough times. Mean breakthrough time in the present analyses (c) is shown as a function of word–shape congruency and shape type. Mean breakthrough time (d) is further shown as a function of word and shape type. All error bars indicate bootstrapped 95% confidence intervals.

breakthrough was affected by low-level visual properties of the stimuli (such as the length of a word or familiarity of the orthography). On the basis of this, as well as the prior finding that a word's frequency does not affect breakthrough times (Heyman & Moors, 2014), Rabagliati and colleagues concluded that there was no evidence for language processing under continuous flash suppression.

If sound symbolism has a replicable effect on breakthrough times, then it presents a strong challenge to Rabagliati et al.'s (2018) conclusion. Figures 1a and 1b display the effect of sound symbolism reported by Heyman and colleagues (2019), who computed a difference score following the analyses of Hung et al. (2017) by subtracting mean incongruent breakthrough

times from mean congruent breakthrough times. Congruent trials refer to a puffy shape containing the word *bubu* or a spiky shape containing the word *kiki*, whereas incongruent trials are a puffy shape containing *kiki* or a spiky shape containing *bubu*. Using the open data and code provided by Heyman and colleagues at <https://osf.io/kwytv/files/>, Rabagliati independently confirmed Heyman and colleagues' finding that there was a significant but small effect of congruency on breakthrough times (mean difference = 0.05 s, 95% confidence interval, or CI = [0.02, 0.08]), $t(178) = 2.75$, $p = .003$, Cohen's $d = 0.05$.

However, Figures 1c and 1d show that the reported effect of congruency does not in fact provide strong evidence for a general sound-symbolism effect.

Participants in Hung et al.'s (2017) and Heyman et al.'s (2019) studies saw only the four stimuli described above, and when the data are broken down by stimulus, a different pattern emerges for each pair of stimuli. There was not a systematic congruency effect; rather, for the puffy shape, seeing the congruent word (*bubu* rather than *kiki*) caused shorter breakthrough times, whereas for the spiky shape it did the reverse. More specifically, no matter whether the shape was puffy or spiky, the label *bubu* always led to faster breakthrough times than the label *kiki*. Mixed-effect regressions confirmed that responses to *bubu* were significantly faster than responses to *kiki* not only for the puffy shape (*bubu*: $M = 3.48$ s, 95% CI = [3.33, 3.63]; *kiki*: $M = 3.81$ s, 95% CI = [3.66, 3.97]), $\beta = 0.34$, $SE = 0.03$, $t(173.4) = 12.6$, $p < .001$, $d = 0.30$, but also for the spiky shape (*bubu*: $M = 3.42$ s, 95% CI = [3.28, 3.56]; *kiki*: $M = 3.65$ s, 95% CI = [3.51, 3.81]), $\beta = 0.23$, $SE = 0.03$, $t(174.6) = 9.1$, $p < .001$, $d = 0.22$ (see the Supplemental Material available online for full analyses and <https://osf.io/tva8j/> for code). These effect sizes were respectively 6 and 4.5 times larger than the omnibus congruence effect size (and it is the slightly larger effect for the puffy shape that caused the original omnibus result).

Another way to look at this is by modeling the data through a mixed-effects regression that treats items as random effects. To this end, Rabagliati regressed breakthrough time on congruency, including a random-effect intercept for each participant, a random-effect intercept for each word, and by-participant and by-word predictors for congruency (see the Supplemental Material for full details). The resulting model showed no significant fixed effect of congruency, $\beta = -0.05$, $SE = 0.1$, $t(1) = 0.57$, $p = .67$. By contrast, without the item random effects, congruency did significantly affect breakthrough, matching the original analysis, $\beta = -0.05$, $SE = 0.02$, $t(166) = 3.16$, $p = .002$. Thus, when item variance is accounted for, the statistical evidence for a generalized sound-symbolism effect seems to dissipate. Note, however, that estimates of random effects are uncertain here because it is hard to draw conclusions about whether an effect varies—or generalizes—across items when only two stimuli are used. An alternative is to incorporate the two items as a fixed effect. That analysis reveals that item interacts with congruency: Incongruent words reliably increase response times for the puffy shape and reliably decrease them for the spiky shape (see the Supplemental Material).

In summary, the congruency effect was directionally inconsistent between the only two pairs of stimuli tested, and the originally reported omnibus effect occurred because the congruency effect for the puffy shape was slightly larger than the reverse congruency

effect for the spiky shape. Thus, it is hard to interpret the results of the sound-symbolism replication. Because there is no neutral condition or baseline to compare with, it is unclear how sound-symbolic congruency affects suppression times. For instance, although the originally reported omnibus effect could in principle be driven by sound symbolism, it could also be driven by other factors, such as idiosyncratic discrepancies in how the individual items affected suppression.

The possibility that idiosyncratic properties of the items drove the reported effects is quite concerning because the structure of the experimental design makes it particularly plausible. In this design, each pair of stimuli was shown 80 times to each participant. As a result, any idiosyncratic effects of the stimuli on suppression times would be repeatedly measured in the resulting data. Since only two pairs of stimuli were used in Heyman et al.'s (2019) study, it would not be possible to tease apart whether these idiosyncrasies caused differences in suppression time or whether sound symbolism caused such differences (a point that also holds for Hung et al.'s, 2017, study). Importantly, there are a number of potential idiosyncratic properties that could cause differences in suppression time. For example, the familiarity of a stimulus is believed to affect suppression (Jiang, Costello, & He, 2007), and it is possible that one of the stimuli may have more closely resembled a prominent brand or logo. Thus, it is quite conceivable that the apparent omnibus effect of sound symbolism could instead be driven by subtle differences between the two pairs of items that were used as stimuli.

Of course, this analysis cannot prove that sound symbolism has no effect on suppression times, and in fact there are other considerations that support its existence. For example, Hung and colleagues' (2017) original article contained a second study, in which subjects first learned two pairs of arbitrary sound-letter associations (counterbalanced across subjects); those stimuli subsequently appeared to elicit a small sound-symbolism effect. Thus, it is possible that idiosyncratic properties of the stimuli are not solely responsible for the observed sound-symbolism effect. Nonetheless, the present reanalysis of the replication data makes a clear case that significant further research is needed in order to draw strong conclusions about unconscious sound symbolism. Specifically, follow-up studies should be required to use a broader range of stimuli in order to establish the generalizability of the effect.

An analogy can perhaps be drawn to a well-established phenomenon such as supraliminal semantic priming. Semantic-priming manipulations have a robust effect on tasks such as lexical decision, but their influence on

response times is still relatively small (i.e., an average speedup of about 25 ms in lexical decision times according to the mega-study by Hutchison et al., 2013). As a result, the effect of semantic priming for any individual item may be inconsistent because of the influence of other factors, such as lexical frequency or word length. Those item-level inconsistencies do not imply that priming does not exist, but we are confident about this assertion only because supraliminal semantic priming has been observed across a wide array of stimuli and contexts. The effect of sound-shape congruency on suppression times is presumably also rather weak (if it exists at all), especially compared with other factors such as pixel density (Heyman & Moors, 2014), and so inconsistent item-level effects are not implausible. However, we have information only about the effect of sound-shape congruency from two of these items, in contrast to semantic priming. Thus, we cannot rule out idiosyncratic properties of the stimuli as a potential alternative explanation unless a replication is conducted with a wider swath of items.

More broadly, the potential impact of idiosyncratic item differences ought to raise worries about the validity and generalizability of other studies of unconscious cognition, as these also often use only a handful of items and rarely incorporate by-item analyses. For instance, prominent prior work on unconscious facial emotion processing (e.g., Gray, Adams, Hedger, Newton, & Garner, 2013; Yang et al., 2007) and unconscious gaze detection (e.g., Stein, Senju, Peelen, & Sterzer, 2011) each used fewer than five sets of stimuli. Clarifying the impact of these concerns, whether through reanalysis or replication with extension, should be an important goal for the field.

The methods and analyses in Heyman and colleagues' (2019) study made sense in the context of a Preregistered Direct Replication, as they closely mimicked the original procedure. However the present finding, that the apparent unconscious sound-symbolism effect is not consistent between the two stimuli used, highlights how replications and preregistered analyses still need careful interpretation. A finding may be reliably replicated, but this still does not guarantee its validity and generalizability. In this case, the exploratory analyses caution against strongly interpreting the findings of the preregistered report: The apparent evidence for unconscious sound symbolism could be artifactual, driven by idiosyncratic differences in how the two item pairs emerged from continuous flash suppression. Thus, and given the state of the literature as a whole, we would argue that language processing without awareness still seems unlikely (Rabagliati et al., 2018). To truly challenge this conclusion, one would

need to demonstrate that the sound-symbolism effect generalizes across a much larger sample of items.

Transparency

Action Editor: D. Stephen Lindsay

Editor: D. Stephen Lindsay

Author Contributions

H. Rabagliati developed the initial critique and conducted the reported analyses. P. Moors and T. Heyman provided critical responses. The final version of the manuscript was drafted and approved for submission by all three authors.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

P. Moors was funded through a postdoctoral fellowship awarded by the Research Foundation – Flanders (Grant No. 12X8218N).

Open Practices

All code used in the present analyses has been made publicly available via OSF and can be accessed at <https://osf.io/tva8j/>. Data used in the analyses were from the study by Heyman, Maerten, Vankrunkelsven, Voorspoels, and Moors (2019) and can be found at <https://osf.io/kqwgc/>. The design and analysis plans for the present study were not preregistered. The complete Open Practices Disclosure for this article can be found at <http://journals.sagepub.com/doi/suppl/10.1177/0956797620949461>. This article has received the badge for Open Materials. More information about the Open Practices badges can be found at <http://www.psychologicalscience.org/publications/badges>.



ORCID iDs

Hugh Rabagliati  <https://orcid.org/0000-0001-9828-5857>

Tom Heyman  <https://orcid.org/0000-0003-0565-441X>

Acknowledgments

The authors thank David Carmel for helpful comments.

Supplemental Material

Additional supporting information can be found at <http://journals.sagepub.com/doi/suppl/10.1177/0956797620949461>

References

- Gray, K. L. H., Adams, W. J., Hedger, N., Newton, K. E., & Garner, M. (2013). Faces and awareness: Low-level, not emotional factors determine perceptual dominance. *Emotion, 13*, 537–544.
- Hassin, R. R. (2013). Yes it can: On the functional abilities of the human unconscious. *Perspectives on Psychological Science, 8*, 195–207. doi:10.1177/1745691612460684

- Heyman, T., Maerten, A.-S., Vankrunkelsven, H., Voorspoels, W., & Moors, P. (2019). Sound-symbolism effects in the absence of awareness: A replication study. *Psychological Science*, *30*, 1638–1647. doi:10.1177/0956797619875482
- Heyman, T., & Moors, P. (2014). Frequent words do not break continuous flash suppression differently from infrequent or nonexistent words: Implications for semantic processing of words in the absence of awareness. *PLOS ONE*, *9*(8), Article e104719. doi:10.1371/journal.pone.0104719
- Hung, S.-M., Styles, S. J., & Hsieh, P.-J. (2017). Can a word sound like a shape before you have seen it? Sound-shape mapping prior to conscious awareness. *Psychological Science*, *28*, 263–275. doi:10.1177/0956797616677313
- Hutchison, K. A., Balota, D. A., Neely, J. H., Cortese, M. J., Cohen-Shikora, E. R., Tse, C.-S., . . . Buchanan, E. (2013). The semantic priming project. *Behavior Research Methods*, *45*, 1099–1114.
- Jiang, Y., Costello, P., & He, S. (2007). Processing of invisible stimuli: Advantage of upright faces and recognizable words in overcoming interocular suppression. *Psychological Science*, *18*, 349–355.
- Mellers, B., Hertwig, R., & Kahneman, D. (2001). Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration. *Psychological Science*, *12*, 269–275. doi:10.1111/1467-9280.00350
- Rabagliati, H., Robertson, A., & Carmel, D. (2018). The importance of awareness for understanding language. *Journal of Experimental Psychology: General*, *147*, 190–208. doi:10.1037/xge0000348
- Sklar, A. Y., Levy, N., Goldstein, A., Mandel, R., Maril, A., & Hassin, R. R. (2012). Reading and doing arithmetic nonconsciously. *Proceedings of the National Academy of Sciences, USA*, *109*, 19614–19619. doi:10.1073/pnas.1211645109
- Stein, T., Senju, A., Peelen, M. V., & Sterzer, P. (2011). Eye contact facilitates awareness of faces during interocular suppression. *Cognition*, *119*, 307–311.
- Yang, E., Zald, D. H., & Blake, R. (2007). Fearful expressions gain preferential access to awareness during continuous flash suppression. *Emotion*, *7*, 882–886. doi:10.1037/1528-3542.7.4.882